B O N N E V I L L E   P O W E R   A D M I N I S T R A T I O N

A Benthic Index of Biotic Integrity (B-IBI) for Mainstem Rivers and Tributaries of the Upper Yakima and Naches River Basins

Final Report 2000 - 2004 | June 2005 | DOE/BP-00006305-2

This Document should be cited as follows:

Gayeski, Nick, "A Benthic Index of Biotic Integrity (B-IBI) for Mainstem Rivers and Tributaries of the Upper Yakima and Naches River Basins", 2000-2004 Final Report, Project No. 200004800, 105 electronic pages, (BPA Report DOE/BP-00006305-2)

Bonneville Power Administration
P.O. Box 3621
Portland, OR 97208

This report was funded by the Bonneville Power Administration (BPA), U.S. Department of Energy, as part of BPA's program to protect, mitigate, and enhance fish and wildlife affected by the development and operation of hydroelectric facilities on the Columbia River and its tributaries. The views in this report are the author's and do not necessarily represent the views of BPA.

A Benthic Index of Biotic Integrity (B-IBI) for Mainstem Rivers and Tributaries of the Upper Yakima and Naches River Basins

Project Final Report
2000 to 2004

Prepared by:
Nick Gayeski, Project Director
Washington Trout
PO Box 402
Duvall, Washington 98019

Contributor:
Kristin Rein,
Portland, Oregon 97212

Prepared For:
U.S. Department of Energy
Bonneville Power Administration
Division of Fish and Wildlife
P.O. Box 3621
Portland, Oregon 97283-3621

Project Number: 200004800
Contract Number: 00006305

JUNE 2005

ACKNOWLEDGEMENTS

The Project Director wishes to thank Kris Rein for her many contribution to the success of this project, from long hours in the field and in the laboratory to the analyses of mainstem river landuse data and numerous discussions about B-IBI and biomonitoring in general. We both wish to acknowledge the following individuals and their parent organization and agencies for their assistance and support: Dr. Jim R. Karr, School of Fisheries and Aquatic Sciences, University of Washington, for advice on sampling design and metric evaluation, and for the use of his Bug Laboratory at the University of Washington for sample processing, invertebrate identification, and sample archiving: Bill McMillan and Ramon Vanden Brulle of Washington Trout for assistance with field collections; Sarah Morley of NOAA Fisheries for assistance with field collections; Cary Kindberg and Joseph Yacker of Washington Trout for preparation of maps and GIS analysis of GAP road density data; Walt Larrick and Al Sherzinger of the Yakima office of the Bureau of Reclamation for interest in the project and for permission to access several sites on and adjacent to Bureau property; Larry Virdon, Boise Cascade Company, Ellensburg office, for access to two sites on Naneum Creek via gates and road owned by Boise Cascade; David Byrnes, project contract officer for Bonneville Power Administration for his interest in the project and his support.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

We evaluated a Benthic Index of Biotic Integrity (B-IBI) for application to tributary streams and mainstem rivers of the Yakima River basin upstream of Union Gap, Washington, including the Naches River basin, as a tool for assessing the biological condition of aquatic habitat, for monitoring activities that may impair aquatic habitat condition, and for monitoring activities intended to improve the condition of aquatic habitat. The B-IBI is a multi-metric index of the biological condition (biotic integrity) of a segment of stream habitat. The integrity score of a stream segment measured at a sampling site is a sum of several independent measures (metrics) of the condition of the aquatic benthic invertebrate community at the site. Benthic invertebrates occupy stream sites for periods as short as weeks or months to as long as three years. Thus the state of the aquatic invertebrate community at any point in time is a reflection of the physical, chemical, and biological conditions experienced by the site at multiple time scales. By using several independent biological metrics, B-IBI can reflect different kinds of disturbance and quantify different levels of impairment to stream condition.

Each component metric in B-IBI measures an aspect of the composition or state of the stream benthic invertebrate community, such as the number of total invertebrate taxa in a sample. The raw metric value (number of taxa, for example) is then assigned to one of three categories and assigned a score of 1, 3, or 5 according to whether the value represents a low, medium, or high value in terms of biological condition. For example, minimally disturbed sites with a high number of total taxa would receive a score of 5 for total taxa richness, while a very disturbed site with a low number of taxa would receive a

1. The B-IBI score for a site is the sum of the metric scores for all metrics chosen to be included in the index.

To develop a B-IBI appropriate to tributary streams and mainstem rivers in the Yakima basin we sampled extensively throughout the Yakima basin upstream of Union Gap for a period of four (4) years, from 2000 to 2003. We sampled a total of 37 tributary and 23 mainstem sites. Of the 37 tributary sites, 24 were sampled in two or more years, resulting in a total of 70 site-years of sampling at tributary sites.[1] Nine of the 23 mainstem sites were sampled in two or more years, resulting in a total of 35 site-years of sampling. In total, we collected samples for 60 different sites (tributary and mainstem combined) and 105 individual site-years (See Table 2).

We validated a set of ten metrics and associated scoring criteria that characterize site conditions in both tributaries and mainstem reaches in the upper Yakima basin. We found that the same set of metrics and scoring criteria are applicable to both tributary streams and mainstem rivers. With ten metrics, site index scores can range from a low of 10 (ten 1's) to a high of 50 (ten 5's). The ten metrics and associated scoring criteria are listed in Table ES1 (Table 9 in the Report). Further details are provided in the Methods section of the Report.

---

[1] A 'site-year' is one sample collection conducted at a single site in a year. For example, if a site was sampled in 2000 and again in 2002, the site would have two site-years of sampling effort.

Table ES1. Metrics scoring criteria for tributary and mainstem sites in Yakima and Naches river basins. Numbers in braces {} in Percent Tolerant metric are criteria for Puget Sound Lowlands B-IBI.

B-IBI Metric	1	3	5
Total Taxa	<=14	(14 - 28)	>=28
Total Mayfly Taxa	<=3.5	(3.5 - 7)	>=7
Total Stonefly Taxa	<=2.7	(2.7 - 5.3)	>=5.3
Total Caddisfly Taxa	<=2.7	(2.7 - 5.3)	>=5.3
Total Clinger Taxa	<=8	(8 - 16)	>=16
Long-Lived Taxa	<=4	(4 - 8)	>=8
Intolerant Taxa	<=2	(2 - 4)	>=4
Percent Predators	<4.5	[4.5 - 9)	>=9
Percent Tolerants	<10 {<27}	[10 - 25] {27 - 44}	>25 {>44}
Percent Dominance	<55	[55 - 75]	>75

Site scores displayed a modest degree of inter-annual variability that permitted the identification of four classes or states of site condition. Site condition classes range from a high of 1 (reference or near-reference) to a low of 4 (severely degraded). The range of index scores associated with each of the four condition classes are given in Table ES2 (Table 13 in the Report).

Table ES2. Recommended Score-Based Site Condition Classification Criteria

B-IBI Site Score Range	Score-based Site Condition Classification
38 – 50	1
28 - 37	2
20 - 27	3
10 - 19	4

The recommended site classification criteria ranked 30% (11 of 37) of the tributary sites as condition-1, 49% (18 of 37) as condition-2, 16% (6 of 37) as condition-3, and 5% (2 of 37) as condition-4.

Mainstem sites displayed less variation in both site scores and site condition. After adopting the final set of metrics and scoring criteria we found only one mainstem site and two site-years that were in condition-1. We found no sites or site-years in condition-4. Of the 23 mainstem sites, 22 (33 of 35 site-years) were in one of the two intermediate conditions. Fourteen (61%) were in condition-2 (moderately impaired) and eight (35%) were in condition-3 (significantly impaired). We believe that this is an accurate reflection of the condition of mainstem sites relative to the range of conditions displayed in tributaries at comparable elevations throughout the study region. That is, we believe that condition-1 sites can occur in all mainstems reaches throughout the study region if anthropogenic conditions are other than they are, rather than that our metrics and/or scoring criteria are too strict or otherwise inappropriate for mainstem reaches. We would have found condition-1 sites if they were present. We also believe that we would have found condition-4 sites if they were present.

Considering the similarity in taxa found in tributary and mainstem sites, the absence of both condition-1 and condition-4 mainstem sites is most likely a reflection of the pervasive and homogenizing influence of river regulation and not a reflection of the distinctiveness of large (5^{th}-order and greater) rivers *per se* relative to smaller rivers and streams. Despite the narrow range of mainstem site conditions, the recommended index will detect departures from current conditions that result from either anthropogenic or natural disturbances. We therefore recommend its use in monitoring and assessment as described in the following sections.

Influence of Road Density and Percent Forest Cover

We evaluated the impact of two broad-scale indices of landuse on B-IBI score/site condition of tributary and mainstem sites: road density and percent forest cover. Road densities were calculated at three scales using Washington State Department of Natural Resources GAP data acquired in 2000: the entire area of the catchment upstream of each site and catchment area within buffers of 200 and 100 meters on each side of the stream upstream of each site. Catchment-wide road density data displayed no correlation with B-IBI score. Road densities within 200- and 100-meter buffers differed little from one another and also displayed no significant correlation with site score. This result may be due to a combination of relatively low road densities in most of the basin, the coarseness of the road data, and the condition of riparian buffers within 100 meters of stream banks. We suspect that the catchment scale may be too large to detect road impacts at the channel unit (riffle) scale. Road density and condition measured at a finer spatial scale relative to sample sites is likely required to assess road impacts.

Land cover was obtained from a Washington Department of Fish and Wildlife Gap Analysis Program (GAP) data set and consisted of land cover polygons derived from a 1991 Landsat TM image with a spatial resolution (minimum polygon size) of 100 hectares (0.01 square kilometers). Percent forest cover did display a clear and statistically significant correlation with site condition at both mainstem and tributary sites. B-IBI score at both tributary and mainstem sites increased as percent forest cover within the catchment increased. Tributaries displayed a stronger positive association with percent forest cover than mainstem sites (regression slope, tributaries: 0.141, R^2 0.28, $p<0.001$; regression slope, mainstems 0.121, R^2 0.151, $p<0.05$).

Recommended Application of the Index for Monitoring and Assessment

Based upon the initial performance of the recommended index in classifying sample site conditions, we believe that the index has value for the purposes for which we conducted the evaluation: assessment of site condition, monitoring of impacts, and monitoring of site response to restoration actions. In view of the interannual variation in scores at a small number of mainstem and tributary sites, however, we believe that site assessment and monitoring will be most reliable if undertaken for periods of two or more consecutive years. This is particularly important at sites that display intermediate levels of disturbance/impairment (condition-2 and condition-3). The data shows that condition-4 (severely impaired) and condition-1 (minimally-disturbed) sites display little interannual variation compared to condition-2 and condition-3 sites. This is consistent with expectations based upon ecological theory, since minimally-disturbed sites should have greater resilience to minor, intermittent disturbance such as flood events than sites with moderately impaired biotic integrity. Severely impaired sites have less potential to achieve high index scores under favorable environmental condition than less severely impacted disturbed sites. Consequently, sites that may be in condition 2 or 3 and those that may be borderline between condition-2 and condition-1 require more than a single year (snapshot) of assessment or monitoring. In general, we expect that three (3) consecutive years of sampling within a maximum period of five (5) years should be required for monitoring purposes, and preferably for assessment purposes as well.

Sampling costs are modest. No more than two (2) hours are required to collect the required number of invertebrate samples and to take basic measurements of stream channel and riparian area condition. Processing of the three replicate samples that are required to be taken at each site, including sorting and identification requires an average of 20 hours of laboratory time. The total cost for these activities per sample site and sampling occasion is approximately $600.00.

Unresolved Issues and Recommended Future Research

It is important to emphasize that the Benthic Index of Biotic Integrity directly measures the biological condition of the benthic food web and indirectly the biophysical condition of a stream/river site. It does not measure the ability of a site or associated stream reach to support fish taxa of interest such as salmonids. For purposes of monitoring the effect of stream-habitat restoration activities on salmonid species, the B-IBI is best employed in conjunction with other measures of stream-habitat condition such as physical conditions in the stream channel and the associated riparian area and indices of fish population condition and individual fish condition.

While the B-IBI has not been fully evaluated in the context of salmonid fish use, individual fish condition, and abundance, this should be an important area of future research. The B-IBI has potential to provide a direct assessment of the quality of a stream reach from a fish's perspective; it measures both the condition of the benthic food web and the condition of specific kinds of invertebrate taxa important to salmonid fish. The specific link between individual metrics employed in a B-IBI and the feeding ecology of salmonid juveniles has not yet been fully taken advantage of. It would be particularly valuable to employ the B-IBI in conjunction with a study of the functional significance of benthic invertebrate taxa for drift feeding salmonids as exemplified in Rader's "functional classification of the drift" (Rader, 1997). Such a study could provide important information on the correlation between B-IBI index scores and the conditions at the stream channel and stream reach scale that are significant from the point of view of salmonid feeding ecology.

1.0 INTRODUCTION

Biological assessment and monitoring tools are needed in the Columbia Basin for purposes of identifying and prioritizing aquatic-habitat sites for preservation and restoration efforts and for monitoring the progress of such efforts. Unlike physical and chemical measures of the quality of running waters which offer only a snap-shot in time of the condition of aquatic habitats, bio-assessment methods are capable of integrating and reflecting cumulative impacts to aquatic habitats both spatially and temporally. Biological organisms such as fish and macroinvertebrates (but also zooplankton and periphyton) reside in the aquatic environment over multiple periods of time, ranging from weeks and months to several years. Both individual species and community characteristics reflect responses to changes in the aquatic environment. Moreover, details of these species- and community-level responses can be partitioned so as to detect both natural and human sources of landscape changes and their consequent impacts upon the biological condition of the aquatic environment.

A good biological assessment and monitoring tool should be relatively easy and cost-effective to employ on a regular basis, easy to understand and communicate to project managers and their superiors, and robust to the normal exigencies of field sampling. The Benthic Index of Biotic Integrity (B-IBI) developed by Dr. James Karr and his students and colleagues in several regions of the Pacific Northwest during the past decade (Karr & Chu 1999, Karr 1998) possesses these features and consequently has great promise as a tool for monitoring and assessing activities throughout the Columbia River Basin that are undertaken with the intention of preserving or recovering the biological integrity of fluvial aquatic habitats. We developed an Index of Biotic Integrity (IBI) based upon the sampling of aquatic benthic invertebrates (a B-IBI) to serve this purpose in the Naches and upper Yakima basins.

The B-IBI is a multi-metric index of the biological condition of fluvial aquatic habitats based upon sampling benthic invertebrates. Multi-metric indices are constructed from individual metrics chosen to reflect both individual taxa and biotic community response to human-caused impacts on the aquatic environment. Metrics are chosen from among four general classes: taxa richness and composition; taxa tolerance/intolerance; feeding ecology; and population attributes such as taxa dominance and abundance. Each metric is scored according to an objective criterion and the scores of each metric then summed to produce a single score that characterizes the biological condition of a site.

Over the past decade B-IBIs have been developed and evaluated in several regions of the Pacific Northwest outside of the mid- and upper-Columbia River Basins by students and colleagues of Dr. James Karr. These IBIs have been shown to be sensitive indicators of stream condition and reliable at identifying the principal kinds of anthropogenic impacts impairing stream condition (Karr and Chu 1999; Kleindl 1995; Patterson 1996; Adams 2001**;** Fore, Karr, and Wisseman 1996). The statistical properties of B-IBIs and of related IBIs developed in the Midwest in the 1980s and early 1990s using fish instead of invertebrate taxa have been evaluated and shown to be robust and amenable to evaluation by standard statistical techniques, including Analysis of Variance, Regression, and Correlation (Fore, Karr, and Conquest 1994; Doberstein, Karr, and Conquest 2000).

B-IBIs are, however, likely to be sensitive to geologic, geomorphological, and climatological conditions at large landscape scales such as the province scale in the Columbia River Basin. Consequently, it is necessary to evaluate component metrics of a B-IBI in provinces in which it has not yet been employed or evaluated.

A properly constructed B-IBI is able to characterize the condition of sampled sites relative to the condition of reference sites and to identify the principal anthropogenic impacts affecting a site and preventing it from achieving the condition of comparable reference sites. Reference sites are either completely undisturbed, pristine stream or river reaches or the least-disturbed sites available for particular stream orders and elevation ranges in the basin of study.

The primary objective of the project was to evaluate a set of metrics that would characterize the general biotic condition of stream habitats at reach and larger spatial scales and that would be sensitive to anthropogenic disturbances known to affect stream and river condition at these scales. We aimed for a set of metrics that would distinguish four(4) or five(5) categories of site condition on an ordinal scale ranging from severely impaired to near-pristine.

A second aim of the project was to identify the types of anthropogenic impacts that were associated with specific ranges of site conditions as characterized by B-IBI site scores.

We examined available landuse information and searched for patterns of association between site condition categories and types and extent of landuse.

A third aim was to initiate an evaluation of the applicability of a multi-metric B-IBI to large (fifth-order and higher) mainstem rivers in the Yakima Basin catchment upstream of Union Gap, including the Naches River catchment. The B-IBI was originally developed for wadable streams and rivers of fourth and smaller (Strahler) stream order, but aquatic invertebrate communities in large rivers also integrate the impacts of anthropogenic disturbance at reach and catchment scales. An assessment method for large rivers comparable to the B-IBI for wadable streams should in principle be possible and would be highly desirable. We made a preliminary effort to develop a B-IBI for the larger mainstems of the study area, and present results separately for tributaries (fourth-order and smaller streams) and mainstem rivers (regulated and/or fifth-order and larger rivers).

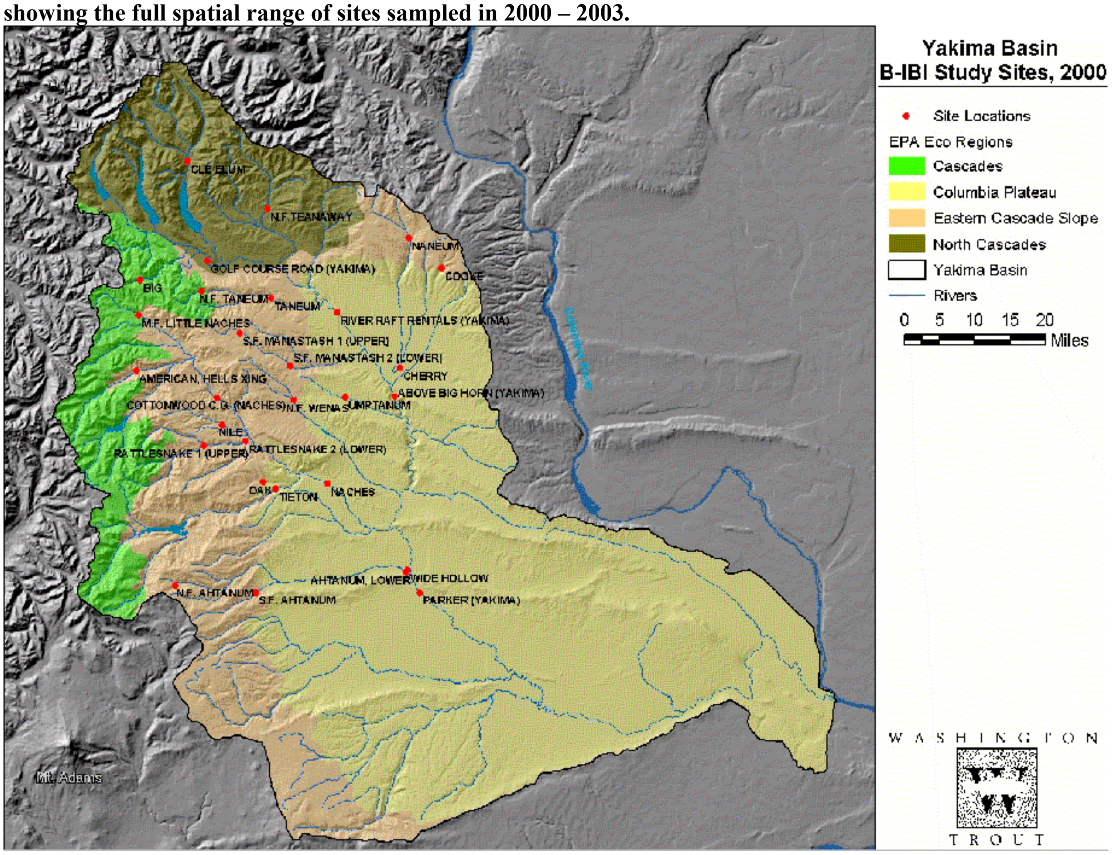
The majority of effort was devoted to extensive field sampling to insure development of a suite of metrics that in aggregate identified four to five categories of stream condition with minimal year-to-year variability. Four months of field work distributed over the four years of the project was devoted to the collection of benthic invertebrate samples over a representative range of stream order, elevation, and stream reach and catchment condition. This insured that we achieved a significant replication in both space and time on which to base our selection of individual metrics. All sampled units were riffles of low to moderate gradient as explained in the Methods section of this Report.

2.0 STUDY AREA

The study area consisted of mainstem river and tributary subbasins of the Yakima River upstream of Union Gap, in central Washington state east of the Cascade crest (Figure 1). For convenience of reference only we divided the Yakima river basin upstream of Union Gap into three mainstem basins: the Naches river basin (Naches), the Yakima river upstream of Roza Dam (upper Yakima), and the Yakima river between Union Gap and Roza Dam, excluding the Naches (lower Yakima). Sampling was conducted at the reach scale, where a reach is defined as a length of stream channel and associated riparian zone of uniform gradient and substrate condition approximately ten to 20 average channel widths in length. Evaluations of the condition of the riparian zone within a reach were normally confined to within ten meters of the edge of each bank of the bankfull channel. After first identifying a stream segment within which to choose a sampling location, we then identified a riffle from which to collect benthic invertebrates according to the procedure described in Methods below. The sample site was then taken to extend upstream from that riffle for a length of ten to 20 channel widths. We frequently refer to a site as a sample reach and generally employ the terms 'site', 'reach', sample site' and 'sample reach' interchangeably.

Sites were chosen to form a representative sample of aquatic stream habitat conditions in the studied basins over a representative range of stream-orders and elevations. (The criteria by which we chose tributary reference sites are discussed separately in the Methods Section below.)

FIGURE 1. Map of the Yakima River Basin with a representative number of sites showing the full spatial range of sites sampled in 2000 – 2003.



Sites sampled in the three mainstem river basins included both mainstem and tributary subbasin sites. We classified a river as a mainstem if it was of fifth or greater stream order or if a major segment of it was regulated and displayed an augmented discharge pattern for part of the year relative to the discharge pattern that would be expected in the absence of regulation. For simplicity's sake we considered the segment of regulated rivers upstream of the influence of reservoir inundation as a mainstem river as well. We classified a river or stream as a tributary if it was of 4^{th} or lower order, unless it was regulated as described above. For example, we classified the Bumping River upstream of

Bumping Reservoir as a mainstem river since it is the mainstem of a river that is impounded and regulated downstream.

Sites sampled ranged in elevation from just below 1000 feet in the immediate vicinity of Union Gap, Washington to over 4000 feet on North Fork Ahtanum and South Fork Manastash Creeks. The downstream-most tributary subbasin sampled was Ahtanum Creek, which enters the mainstem Yakima River at Union Gap. The upstream-most site sampled was the mainstem of the Yakima River upstream of Roza Dam at Crystal Springs.

Tributary subbasins in the Naches basin included the Tieton, Bumping, Little Naches and American Rivers, and Rattlesnake Creek. Tributary subbasins sampled in the upper Yakima River Basin included Naneum and Cooke Creeks in the Colockum Hills above the Kittitas Valley, Manastash Creek, Taneum Creek, Big Creek, Cabin Creek and the Teanaway and Cle Elum river subbasins. The distribution of sampled sites by elevation and major stream type is listed in Table 1.

Table 1. Number of site-years sampled by elevation range and stream type for sample years 2000 - 2003

Elevation (feet)	Mainstem	Tributary	Total
1000 - 1500	13	15	28
1500 - 2000	8	8	16
2000 - 2500	10	15	25
2500 - 3000	2	13	15

3000 - 3500	1	8	9
3500 - 4000	1	7	8
Above 4000	0	4	4
Total Site-Years	35	70	105

2.0.1 Timing and Frequency of Sampling

All sampling was conducted in late summer during tributary baseflow, as recommended by Dr. James Karr for B-IBIs in the Pacific Northwest (Karr 1998). We intended to sample approximately 30 sites per year during the first three years of the project (2000 to 2002) within a 30-day window between mid-August and mid-September. Sampling during the fourth year of the project (2003) was devoted to sampling sites chosen based on the data for the first three years, to enhance replication at particular locations or kinds of sites and to add to the number of samples at sites that exhibited unexpected variability. Since the upper Yakima and Cle Elum rivers have augmented summer flows through August and the Naches, Tieton, and Bumping rivers have augmented flow beginning in early September due to the Flip-Flop regime, we sample sites on the Tieton, Naches, and Bumping mainstems and on the mainstem of the Yakima downstream of the Naches river during the third and fourth weeks of August and on the mainstem of the upper Yakima during the second week of September.

Frequency of sampling is given in terms of site-years, where a site-year is defined as one site sampled in one year. We sampled 32 sites in 2000, 28 in 2001, 33 in 2002, and 12 in 2004, for a total of 105 site-years. These 105 site-years were distributed over 60 distinct sites, divided between 23 distinct mainstem sites, and 37 distinct tributary sites. Table 2 lists the distribution of sampled sites by major stream type and number of site-years of sampling and Table 3 lists the distribution of sampled sites by mainstem and tributary subbasin.

Table 2. Sample frequency (number of site-years) by Stream type for sample years 2000 - 2004

Frequency of Sampling	Mainstem	Tributary	Total
One Site-Year	14	13	27
Two Site-Years	7	16	23
Three Site-Years	1	7	8
Four Site-Years	1	1	2
Total Multiple-Year Sites	9	24	33
Total Site-Years	35	70	105

Table 3. Number of site-years sampled by Subbasin. Letters in parentheses designate the mainstem subbasin in which each tributary subbasin is located. Mainstem segment abbreviations: Lower Yakima River (downstream of Roza Dam), LY; Upper Yakima River (upstream of Roza Dam), UY; Naches River (N).

Subbasin or Mainstem Segment	Site-Years Sampled
Lower Yakima Mainstem	5
Upper Yakima Mainstem	12
Naches River Mainstem	9* (8 mainstem, 1 tributary springbrook)
Ahtanum Creek (LY)	8
Wide Hollow Creek (LY)	2

Moxee Drain	1
Wenas Creek (LY)	5
Umptanum Creek (UY)	1
Manastash Creek (UY)	4
Cooke Creek (UY)	3
Naneum Creek (UY)	4
Cherry Creek (UY)	2
Taneum Creek (UY)	10
Teanaway River (UY)	4
Cle Elum River (UY)	3
Big Creek (UY)	4
Cabin Creek (UY)	3
Cowiche Creek (N)	3
Tieton River (N)	5
Rattlesnake Creek (N)	5
Nile Creek (N)	2
Bumping River (N)	4
Little Naches River (N)	2
American River (N)	4
Total	105

2.0.2 Weather Patterns During the Project Period

The sampling period coincided fortuitously with a range of environmental conditions typical of the study area over the past two decades. Table 4 shows summary rainfall and snow depth data for Stampede Pass, Washington near the head of the basin and rainfall data for the city of Yakima, Washington.

Table 4. Precipitation indices for project sampling period, 1999 – 2003. Maximum and average snow depths are for the period October 1 through May 31. Numbers are for precipitation and snow depth in inches at Stampede Pass, Washington and (in parentheses) for Yakima, Washington.

Year	October 1 – July 31	October 1 – May 31	June & July	Max. Snow Depth	Ave. Snow Depth
1999-2000	52.95 (3.48)	49.18 (3.43)	3.77 (0.05)	105	64
2000-2001	45.62 (4.57)	39.78 (3.49)	5.84 (1.08)	70	41.6
2001-2002	62.81 (7.13)	58.28 (6.30)	4.53 (0.83)	156	80.2
2003-2004	59.21 (8.33)	57.66 (8.33)	1.55 (0.0)	111	51.1

Relatively low-water conditions prevailed during the summers of 2000 and 2001. In 2001 very low rainfall amounts through May 31 were reflected in the lowest maximum and average snow depths recorded at Stampede Pass during the four-year period. Summer rainfall through June and July was lowest in 2000 and 2003 and wettest in 2001. 2001 - 2002 was significantly wetter than all other years based upon total rainfall through both May 31 and July 31 and maximum and average snow depths. Based on this data, we believe that our evaluation of metrics for the B-IBI will be robust to normal environmental variation as indexed by precipitation.

3.0 METHODS

3.0.1 Sample Collection and Processing

We followed a standard sampling protocol employed by Dr. James Karr and colleagues at the University of Washington for collecting benthic invertebrate samples in wadable streams for purposes of assessing benthic index of biotic integrity (B-IBI). All samples were collected using a Surber sampler built to Dr. Karr's specifications by Research Nets of Redmond, Washington. The frame of the sampler covered an area of stream bottom 0.09 square meters (30 cm. x 30 cm.). Netting was 500 micron Nytex and terminated in a 4-inch diameter removable PVC cod-end cup. The cod-end cup was vented along the sides and screened with 500 micron Nytex to prevent backwashing when the sampler was placed in high-velocity riffles.

We worked in pairs or, occasionally, in groups of three. A total of five persons participated in collecting the samples in the field during the four years, including three (the Project Director, Nick Gayeski, Kris Rein, and Sarah Morley) who had extensive experience with the sampling protocol. Two other persons (Bill McMillan, and Ramon Vanden Brulle) sampled as part of a two or three person team with one of the other three during the first three years and became highly proficient in the conduct of the sampling protocol. BM and RVB conducted all sampling in the fourth year (2003). This insured a consistent sampling effort within and between sites, and within and between years.

Within each reach chosen for sampling the sampling crew used best judgment to choose the best riffle-habitat unit available. Riffles were chosen on the basis of gradient, substrate condition and heterogeneity of gravel and cobble sizes, with a minimum of fines and a minimum of large cobbles or boulders. In both tributary and mainstem sites, the majority of riffles were sampled at depths of 30 to 60 centimeters and as close to the thalweg as this sampling depth permitted. On rare occasions the depth sampled at mainstem sites was greater than 60 centimeters and the depth sampled at tributary sites was less than 30 centimeters. Within each riffle three replicate samples were taken approximately five feet apart from one another, starting at the downstream end and taking successive replicates upstream of the preceding replicate.

Each replicate sample was taken by first placing the bottom frame of the sampler firmly on the substrate with the mouth of the sampler facing directly into the flow of the current and allowing the net and cod-end to straighten out downstream. Large cobbles lying on the surface of the bottom within the frame of the sampler were then systematically rubbed by hand for a period of 30 seconds to dislodge insects and insect casings attached to the surface. This was immediately followed by disturbing the finer substrate to a depth of 10 centimeters with an asparagus knife for an additional 30 seconds. The opening of the sampler was then lifted above the surface of the water, and insects and debris clinging to the inside of the netting rinsed into the cod-end piece (the rinsing vessel was screened at 500 microns to avoid adding invertebrates or other flotsam to the sample replicate). The sampler was then carried to a working area on the stream bank and the contents of the cod-end piece carefully transferred to a white bucket for additional field processing.

Replicates were processed in the field to minimize the amount of plant material, detritus and fines that had to be picked through in the laboratory. All insects together with plant material, detritus, and fines not discarded were transferred into one or occasionally two 120 ml screw-top benthic jars, filled with ethyl alcohol and sealed. Several turns of electrical tape were placed around the lid to insure that no leakage would occur, and the jar then placed in a quart-sized Zip-Loc Freezer bag. Waterproof paper labels identifying the site, the date the sample was collected and the replicate number (1 – 3) were placed in the jar before it was sealed and in the Zip-Loc bag and the information entered in a field notebook on-site to insure accurate identification of the sample in the laboratory. All samples were processed and insects identified in Dr. Karr's Bug Laboratory at the University of Washington by one of us (KR) as described later in this section.

At the majority of sites we also collected quantitative and qualitative measurements of channel habitat and riparian condition within the reach during the initial visit to the site, after first collecting the three replicate invertebrate samples. We measured bankfull channel width, measured or estimated reach gradient, conducted a Wolman pebble count (n = 100 pebbles), estimated percent canopy (trees and shrubs > 5-meters high), percent understory (trees and shrubs < 5-meters high), and percent groundcover within ten meters of the bankfull edge of both banks, noted bank condition (hardening and erosion), channel sinuosity, and the presence or absence of woody debris in the channel and adjacent to the bankfull edge. This information was used to help characterize the qualitative condition of the site at the reach scale as described below.

3.0.2 Laboratory Procedures

In the lab, each sample was processed separately. Benthic invertebrates were removed under a dissecting microscope, identified to lowest practical taxonomic level (usually genus for benthic insects; class or order for other benthic invertebrates) and counted. Adults, pupae and non-benthic invertebrates were discarded. We counted all samples fully during the first year, but the average abundance was higher than expected. So due to constraints in time and budget, in subsequent years we employed a subsampling methodology. Each sample was emptied entirely and distributed evenly into a tray with gridded squares drawn on the bottom. Squares were selected randomly and all invertebrates were counted in each square until 1) the entire sample was counted or 2) at least 700 individuals were counted. If we reached 700 individuals while counting a square, the remaining individuals within that square were also counted so subsampling often gave abundances greater than 700.

3.0.3 Metrics Evaluation

The Benthic Index of Biotic Integrity is an additive, multi-metric index that is the sum of the scores of several component metrics. The individual component metrics themselves directly reflect biological features of the sampled benthic invertebrate community. Metrics are chosen to reflect one or more of four basic features of benthic invertebrate community structure: (1) taxa richness and composition, (2) tolerance and intolerance with respect to substrate disturbance, fine sediment input, and toxic chemical inputs, (3) feeding ecology and habits, and (4) population attributes. Table 5 lists the metrics for each of these categories that were chosen for the final Index in our study.

The full range of raw values of each of the metrics observed in the entire data set is divided into three mutually exclusive and exhaustive sub-ranges and each sub-range is then assigned a metric value of 1, 3, or 5, depending upon whether the sub-range is

Table 5. Hypothesized response to human disturbance of invertebrate assembly attributes (after Karr & Chu 1999, Table 7, page 77)

METRIC	PREDICTED RESPONSE TO HUMAN DISTURBANCE
Taxa Richness & Composition:	
Total Number of Taxa	DECREASE
Total Number of Ephemeroptera Taxa	DECREASE
Total Number of Plecoptera Taxa	DECREASE
Total Number of Trichoptera Taxa	DECREASE
Total Number of Long-Lived Taxa	DECREASE
Tolerance/Intolerance	
Total Number of Intolerant Taxa	DECREASE
Percent tolerant Taxa	INCREASE
Feeding Ecology and Habits	
Total Number of Clinger Taxa	DECREASE
Percent Predator Taxa	DECREASE
Population Attributes	
Percent Dominance (3 most abundant taxa)	INCREASE

indicative of a severely disturbed, moderately disturbed, or minimally disturbed condition, respectively. An appropriate suite of ten to twelve metrics when summed provides an informative index of relative biological condition at a site (Karr & Chu 1999; Karr 1998; Fore, Karr, and Wissman 1996).

The invertebrate data from the entire four years of sampling were examined in conjunction with the descriptions of the qualitative and quantitative condition of the reach to arrive at a final set of metrics and scoring criteria. The evaluation and selection of metrics and metric-scoring criteria is an iterative process involving the repeated comparing of site-condition, as described by quantitative measures and by qualitative characterization, with site-condition as characterized by the individual metric values.

Specifically, we initially chose a set of mainstem and tributary sites that we expected would encompass a representative array of site conditions over a range of elevations, landscapes, and landuses representative of the Naches river basin and the Yakima river basin upstream of Union Gap. Prior to analyzing any of the invertebrate samples that were collected we characterized site condition qualitatively on a scale of 1 to 4, as described in the next subsection. After processing and summarizing the invertebrate samples as explained below, we partitioned the observed range of the raw values of each candidate metric into three mutually exclusive sub-ranges and assigned a score of 1, 3, or 5 to each sub-range according to whether the sub-range indicated severely, moderately, or minimally disturbed condition, respectively, for the particular biotic attribute. The individual metric scores (1, 3, or 5) for each of the several metrics being tested (approximately ten) were added to arrive at the provisional index score for the condition of each site. The resulting array of index site scores were then inspected to see if the array appeared to display a distribution closely approximating the distribution of qualitative characterization of site conditions. For example, with ten metrics, the highest score a site can obtain is a 50 (10 x 5) and the lowest a 10 (10 x 1). If we had characterized ten of 40 sites as condition 1, ten as condition 2, ten as condition 3, and ten as condition 4, we might expect approximately equal numbers of sites scoring between 10 and 20, 20 and 30, 30 and 40, and above 40.

If the two distributions appeared to us to be close we next examined the particular pairings of index score and qualitative site characterization. Using the example above, we would expect most or all of the index scores between 10 and 20 to correspond to sites characterized as condition 4, and the most or all of the scores greater than 40 to correspond to condition-1 sites.

If discrepancies occurred but were not too great we attempted to resolve them in one of two ways. First, we retained the individual metrics but re-examined the way we divided the range of raw metric values into sub-ranges, re-calculated the index score and again compared the rankings of site condition indicated by the index score to the qualitative characterization that we initially gave the site. If such scoring adjustments seemed to improve the overall match we tentatively retained the revised scoring criteria. Second, if minor scoring adjustments did not appear to make a noticeable difference in the pairings but we believed that the metric scores were sound (i.e., the metrics themselves were broadly responding to site condition in an expected fashion), we considered revising our initial qualitative characterization of the conditions of individual sites. We used the biology as reported by the tentative metrics to temper and correct out initial impressions of what the condition of a site was. We say more about the details of this process in the subsection of Results , “Correspondence Between Qualitative Site Classification and B-IBI Score”. If the discrepancies were large, the metric was discarded.

3.0.4 Qualitative Site Characterization

In order to start the process of evaluating the candidate metrics described above, we made preliminary qualitative characterizations of each stream habitat site sampled. Sites were ranked on a four-point ordinal scale of 1 to 4, where a 1 designated a site judged to be minimally-disturbed by anthropogenic impacts (a reference site) and a 4 designated a severely impacted site. Sites were classified as 2 or 3 based upon their departure from a minimally-disturbed condition, as described below. A site classified as a 2 as opposed to a 3 was judged to be significantly closer to attaining the condition of a 1 than a site judged a 3.

Mainstem (regulated) and tributary sites were characterized separately from one another using different qualitative and semi-quantitative attributes as the basis for classification due to the significant differences in the kinds and spatial scales of disturbance to which each type is subject. The criteria employed in assigning one of the four designations to a site is described separately for tributary and mainstem sites in the following subsections.

3.0.5 Tributary Site Characterizations

In evaluating each tributary site we considered the condition of the channel and the riparian corridor of the sample reach. Reach-scale observations were then supplemented with road density data for tributary subbasins contained in Washington Department of Natural Resources GAP-Analysis databases. During the initial survey at each site we evaluated the riparian area within approximately ten meters of the bankfull edge from the riffle in which samples were collected upstream approximately 20 channel widths. We visually estimated the percent composition of canopy (trees and shrubs greater than five meters in height), understory (trees and shrubs 0.5 to five meters in height), and groundcover (grasses and woody vegetation less than 0.5 meters).

At the time of the initial site survey at most sites we also took a Wolman pebble count (100 random pebbles) in the sampled riffle after taking the benthic invertebrate sample, and also noted any signs of fine sedimentation in excess of the level we expected to observe if the site were undisturbed. Pebble counts were later examined to identify signs of excessive bedload movement (coarsening), excessive fine sediment loading or absence of competent flow (fining) indicative of altered hydrologic regime and/or fine sediment input in the catchment upstream of the site.

We judged a stream reach to be “minimally-disturbed” (condition 1) if the total area of the riparian corridor within 10 meters of the bankfull edge on both banks consisted of a minimum of 70% canopy and understory, if less than 10% of the total length of the stream reach (both banks combined) had hardened or eroding banks, if stream substrate appeared to possess a heterogeneous distribution of pebble sizes free of excessive fines and boulders (>256 mm B-axis diameter), and if the channel appeared to possess a normal degree of sinuosity. We did not require large woody debris to be present in or over the bankfull channel or at its edge unless the catchment was expected to be heavily wooded under minimally-disturbed conditions and woody debris was expected to contribute significantly to channel structure and complexity. However, we did count the presence of woody debris recruited locally from the riparian corridor as contributing toward classifying a site as minimally-disturbed, particularly in the more arid tributaries where the riparian corridor was dominated by shrubs.

We judged a stream reach to be severely disturbed (condition 4) if it was excessively straightened or bank-hardened, if a large percentage of either bank within the reach was eroding, and if the substrate possessed excessive (>25%) fines (< 2mm diameter) or algal growth or other indication of excessive nutrient or other input from agricultural run-off or grazing. We also judged a stream to be severely disturbed if it lacked a moderately-developed riparian corridor with more than 25% of the total area in canopy and understory combined or if the substrate was composed of more than 30% boulders.

Assigning a site to condition 2 or 3 involved first judging the site to be neither minimally- nor severely disturbed and then evaluating the extent to which it appeared to depart from condition 1. A site was classified as condition 2 if channel sinuosity appeared to be consistent with the sinuosity expected for a minimally-disturbed stream reach of similar stream order, gradient and confinement, but one or two of the other primary attributes (riparian composition and condition, substrate condition, bank stability/erosion) departed moderately from the condition required for classification as minimally-disturbed. Proximity of roads and the severity of road-runoff and dust was a major factor in differentiating condition 1 and 2 sites in the majority of instances. A site was classified as in condition 3 if one of the following was present and other attributes departed no more than condition 2 sites: reduced channel sinuosity , extensive band hardening or bank erosion, degraded and/or simplified riparian corridor lacking in canopy and understory.

3.0.6 Choice of Tributary Reference (Condition 1) Sites

Tributary reference (minimally-disturbed) sites were chosen from among all subbasins surveyed. We required candidate reference sites to fit our qualitative site characterizations. We endeavored to choose reference sites from subbasins that were similar in basic geography, vegetation, and elevation to those in which typical landuses occurred. Ideally we would have reference sites within each 500-foot elevation interval between 1000 and 4000 feet in the study area and in the same proportions as all other sites.

As is common with referenced-based biomonitoring protocols, we were not able to find reference sites in the lowest elevation ranges due to the prevalent impact of human disturbance at lower elevations. We identified 12 reference sites ranging in elevation from 2200 to 4700 feet. Two of the 12 were at elevations above 4000 feet, two were between 3500 and 4000 feet, four were between 3000 and 3500 feet, two were between 2500 and 3000 feet, and two were between 2000 and 2500 feet. Two sites were in the

Lower Yakima basin, in the Ahtanum Subbasin. Three sites were in the Naches basin, and seven were in the Upper Yakima basin. Despite the absence of reference sites at elevations below 2000 feet, we are confident that the spatial distribution of the 12 reference sites provides a broad representation of the composition of benthic invertebrate communities in minimally-disturbed streams throughout the study area sufficient to serve the purposes required of reference sites in developing a robust B-IBI.

A representative list of tributary sites, initial condition classification, elevation, and major subbasin is given in Table 6. Table 7 lists all 12 reference sites, elevations, mainstem subbasin, and B-IBI score. The complete list of tributary sites and B-IBI scores at each site averaged over all years of sampling is given in Table 10.

Table 6. Representative Tributary Sites and Qualitative Site Condition Classification

Tributary Site	Site Condition	Elevation	Major Subbasin
South Fork Ahtanum	1	2500	Ahtanum
Big Creek 1	1	3600	Upper Yakima
Big Creek 2	1	2200	Upper Yakima
Middle Fork Little Naches	1	3100	Naches
Upper Rattlesnake Creek	1	2700	Naches
Cooke Creek	1	3500	Upper Yakima
Naneum Creek 1	1	3500	Upper Yakima
North Fork Taneum Creek 1	1	3700	Upper Yakima
Oak Creek	2	2500	Naches

Nile Creek	2	2400	Naches
Lower S. F. Manastash Creek	2	2800	Upper Yakima
Taneum Creek 1	2	2700	Upper Yakima
Taneum Creek 2	2	2000	Upper Yakima
North Fork Wenas Creek	2	2500	Upper Yakima
Lower Rattlesnake Creek	2	2000	Naches
Little Naches	2	2600	Naches
Cowiche Creek	3	1500	Naches
Ahtanum Creek	3	1000	Ahtanum
Cabin Creek	3	2300	Upper Yakima
Eschback Spring Brook	3	1300	Naches
Taneum Creek 3	3	1900	Upper Yakima
Naneum Creek 2	3	2800	Upper Yakima
Taneum Creek 4	3	2000	Upper Yakima
Cherry Creek	4	1500	Upper Yakima
Wenas Creek	4	1200	Upper Yakima
Wide Hollow Creek	4	1000	Wide Hollow
Moxee Drain	4	1000	Moxee

Table 7. Tributary Reference (Condition 1) Sites, Elevations and B-IBI Scores. Numbers in parentheses are the number of years each site was sampled. Scores for sites with multiple years of sampling are average scores.

Site-Code	Site detail	Elevation	Mainstem Basin	B-IBI Score
AHT1	North Fork Ahtanum (2)	4700	Lower Yakima	36

AHT2	South Fork Ahtanum (3)	2500	Lower Yakima	43
AME2	American above Hell's Crossing (2)	3300	Naches	31
BIG1	Big Creek (3)	3600	Upper Yakima	35
BIG2	Big Creek 2 (1)	2200	Upper Yakima	38
COO1	Cooke Creek (3)	3500	Upper Yakima	43
LNA1	M. F. Little Naches (1)	3100	Naches	36
MAN1	S.F. Manastash, upper (2)	4300	Upper Yakima	37
NAN1	Naneum 1 (2)	3500	Upper Yakima	38
NAN3	Naneum 3 (1)	2800	Upper Yakima	42
NTA1	North Fork Taneum 1 (2)	3700	Upper Yakima	40
RAT1	Rattlesnake, upper (4)	2700	Naches	36

3.0.7 Mainstem Site Characterizations

Mainstem sites presented a significantly different problem than tributaries in determining qualitative classification of site condition. With the exception of the upper Bumping River and upper Cle Elum River (at Salmon La Sac) all mainstem sites were on regulated rivers. Only the upper Bumping qualified as a minimally-disturbed (condition 1) site, but due to elevation (3700 feet), geology, and relatively small size it could not provide a representative standard against which to judge the remaining mainstem sites. Because of the pervasive influence of regulation on river condition downstream of the points of regulation, no mainstem reach appeared to us to qualify as a minimally-disturbed site that would facilitate identifying a meaningful range of site conditions (condition 1). Neither did we find a mainstem site upstream of Sunnyside Dam at Parker that appeared to us to qualify as severely disturbed (condition 4).

Had we followed the approach employed in classifying tributary sites, we would have been forced to choose one or more of the best available mainstem sites and designate it (them) as condition 1 sites and all remaining sites as condition 2. We judged that such an approach would deprive the exercise of developing a B-IBI for mainstem sites of much of its purpose and usefulness. Consequently, we chose to evaluate all other mainstem sites against an estimate of what conditions of the principal mainstem sites would be in the absence of regulation. We achieved this by combining information obtained from our sampling of unregulated tributary sites with information from sites on the upper Naches River where the influences of regulation are minimal.

Characterizations of mainstem site conditions were based primarily upon evaluation of channel sinuosity, riparian corridor and floodplain condition relative to valley width, and channel bed-sediment characteristics. Regulated rivers of fifth and larger stream order are typically characterized by reduced flood peaks, altered temporal pattern of the hydrograph, altered timing and prolonged duration of bankfull and near-bankfull flows, altered timing and prolonged duration of baseflow, altered thermal pattern, and coarsened bed sediments. In addition, flows often fluctuate between near-bankfull and near-baseflow more frequently than unregulated rivers and do so asynchronously with environmental conditions such as rainfall that are normally correlated with stage fluctuations in unregulated rivers. Prolonged duration of near-bankfull flows during summer months is generally associated with a simplified channel morphology and an altered and simplified riparian community, frequently dominated by invasive grasses (Stanford and Ward 1992, 1995; Stanford et al 1996).

A site was judged to be in condition 2 if it possessed a riparian corridor dominated by tree canopy and shrub understory, if the channel within the reach was relatively sinuous, if a majority of the bed sediment was widely distributed (greater than 80%) within the range of 8- and 196-mm B-axis diameter, and if less than 20% of both banks combined were hardened or eroding. A site was judged to be in condition 3 if the riparian corridor within the reach contained less than 50% canopy and understory, if it lacked sinuosity, if more than 20% (but less than 50%) of both banks combined were hardened or eroding, or if particles smaller than 8-mm or larger than 196-mm together composed more than 20% (but less than 50%) of the bed sediment size distribution.

Sites were also categorized as condition 3 if the size distribution of bed sediments was strongly repulsed with 50% or more of the size distribution lying within adjacent phi (log2-mm) classes (e.g., within 32- and 64-mm or within 64- and 128-mm B-axis diameter) or if gravels and cobbles were embedded in fines or covered by large amounts of filamentous algae or detritus indicative of agricultural or urban run-off and resultant eutrophication..

We classified a site as condition 4 if it showed evidence of extreme eutrophication, if the channel was extremely straightened, if more than 50% of banks area was either hardened or eroding, if sediments smaller than 8-mm or larger than 196-mm individually constituted more than 30% of the size distribution or together composed more than 50%

of the distribution, or if one phi size class contained 50% or more of the bed sediment size distribution.

A representative list of mainstem sites, initial condition classification, elevation, and major subbasin is given in Table 8.

Table 8. Representative Mainstem Sites and Qualitative Site Condition Classification (Lower Yakima River: downstream of Roza Dam). Upper Yakima River: upstream of Roza Dam)

Mainstem Site	Site Condition	Elevation	Major Subbasin
Yakima R. at WDFW Game Access	2	2100	Upper Yakima
Cle Elum R. at Roslyn (below Lake)	3	2000	Upper Yakima
Yakima R. at River Raft Rentals	3	1700	Upper Yakima
Yakima R. at Ringer Road	3	1400	Upper Yakima
Yakima R. below mouth of Wilson Cr.	3	1400	Upper Yakima
Yakima River above Big Horn	2	1400	Upper Yakima
Yakima R. upstream Hwy. 24. Bridge	2	1000	Lower Yakima
Yakima R. Downstream Hwy. 24 Br.	3	1000	Upper Yakima
Naches R. at Wapatox Canal	3	1600	Naches
Naches R. at Cottonwood Campground	2	2200	Naches
Bumping R. at Cedar Springs C.G.	2	2800	Naches
Bumping R. below Bumping Lake	3	3400	Naches
Tieton R.	3	1900	Naches

4.0 RESULTS

4.1 Tributary Sites

4.1.1 Metrics and Metric Scoring Criteria for Tributaries

Analyses supported the choice of the same ten (10) metrics as those used in the B-IBI for the Puget Sound Lowlands (Karr 1998, Table 20.3). Scoring criteria were practically identical to those for Puget Sound, with the exception of Percent Tolerants for which our scoring criteria were more stringent. The metrics are listed below and the metrics together with their scoring criteria are presented in Table 9.

- Total taxa: average number of total taxa within a site;
- Ephemeroptera taxa: average number of mayfly taxa within a site;
- Plecoptera taxa: average number of stonefly taxa within a site;
- Trichoptera taxa: average number of caddisfly taxa within a site;
- Clinger taxa: average number of clinger taxa within a site;
- Long-lived taxa: cumulative number of long-lived taxa within a site;
- Intolerant taxa: cumulative number of intolerant taxa within a site;
- Percent tolerant individuals: average percentage of tolerant individuals within a site;
- Percent Predator individuals: average percentage of predator individuals within a site; and
- Percentage dominance: average percentage of individuals in the three most abundant taxa within a site.

Averages are taken over the three replicate samples at each site. Cumulative numbers are the total numbers of distinct long-lived or intolerant taxa in the three replicates taken at a site.

Table 9. Metrics scoring criteria for tributary sites in Yakima and Naches river basins. Numbers in braces {} in Percent Tolerant metric are criteria for Puget Sound Lowlands B-IBI.

B-IBI Metric	1	3	5
Total Taxa	<=14	(14 - 28)	>=28
Total Mayfly Taxa	<=3.5	(3.5 - 7)	>=7
Total Stonefly Taxa	<=2.7	(2.7 - 5.3)	>=5.3
Total Caddisfly Taxa	<=2.7	(2.7 - 5.3)	>=5.3
Total Clinger Taxa	<=8	(8 - 16)	>=16
Long-Lived Taxa	<=4	(4 - 8)	>=8
Intolerant Taxa	<=2	(2 - 4)	>=4
Percent Predators	<4.5	[4.5 - 9)	>=9
Percent Tolerants	<10 {<27}	[10 - 25] {27 - 44}	>25 {>44}
Percent Dominance	<55	[55 - 75]	>75

4.1.2 Metrics Performance: Tributaries

The average tributary site B-IBI scores and sample standard deviations for sites sampled in multiple years are listed in Table 10. The averages and sample standard deviations for all sites within each disturbance category and for the aggregate tributary data set are listed in Table 11. The relationship between disturbance category and average site score is plotted in Figure 2.

Table 11 and Figure 2 indicate that the set of metrics and associated scoring criteria result in a multimetric index (B-IBI) that is moderately successful in discriminating among the four qualitative site conditions. The sample standard deviations are larger than desirable for the purposes of constructing 95% confidence intervals (CIs) narrow enough to distinguish four or five categories of site condition, which would require uniform standard deviation across all site conditions no greater than two. However, we believe that at this stage in the evaluation of a B-IBI for monitoring and assessment purposes in the Yakima-Naches basins such a standard is unreasonably high.

Table 11 shows that the standard deviation of most site scores is less than six. In addition, sites classified as minimally-disturbed (reference) and as severely-disturbed (category 4) have lower standard deviations than moderately-disturbed sites. This is as expected. Minimally-disturbed sites should be resilient to normal environmental variation and disturbance as a result of a greater degree of biotic community structure and/or the greater integrity and complexity of physical stream and riparian habitat. Severely-disturbed sites should display lower variance in B-IBI score because they lack the capacity to score much above their mean score level due to a restricted number of taxa, fewer niches, and a simplified community structure.

The moderately greater variability of condition 2 and 3 streams results from three factors. One is that these are the most difficult sites to classify qualitatively in large part because they exhibit a variety of minor to moderate impacts that produce the impression that they are not in pristine or minimally-disturbed condition. A second is that these sites can be expected to have a more heterogeneous pattern of response to normal environmental perturbations than condition 1 and 4 sites, for the same reasons that condition 1 and 4 sites do not. We should expect condition 2 and 3 sites to exhibit a greater range of responses to normal environmental variation.

A third factor is more directly important. The greater variability is largely the result of between-year variation in B-IBI score at two condition-2 sites and one condition-3 site, each of which was sampled in two years. The sites are listed in Table 10: Oak Creek (OAK1);North Fork Wenas Creek (WEN1); and the North Fork of the Teanaway River (TEA1). In each case the two site scores differed by more than ten points (scores at two sites were 12 points apart; the scores at the third site were 14 points apart).

Table 10 Tributary sites, years of sampling, elevation, site condition, and average B-IBI score. Sites ordered by disturbance condition and by elevation within disturbance condition.

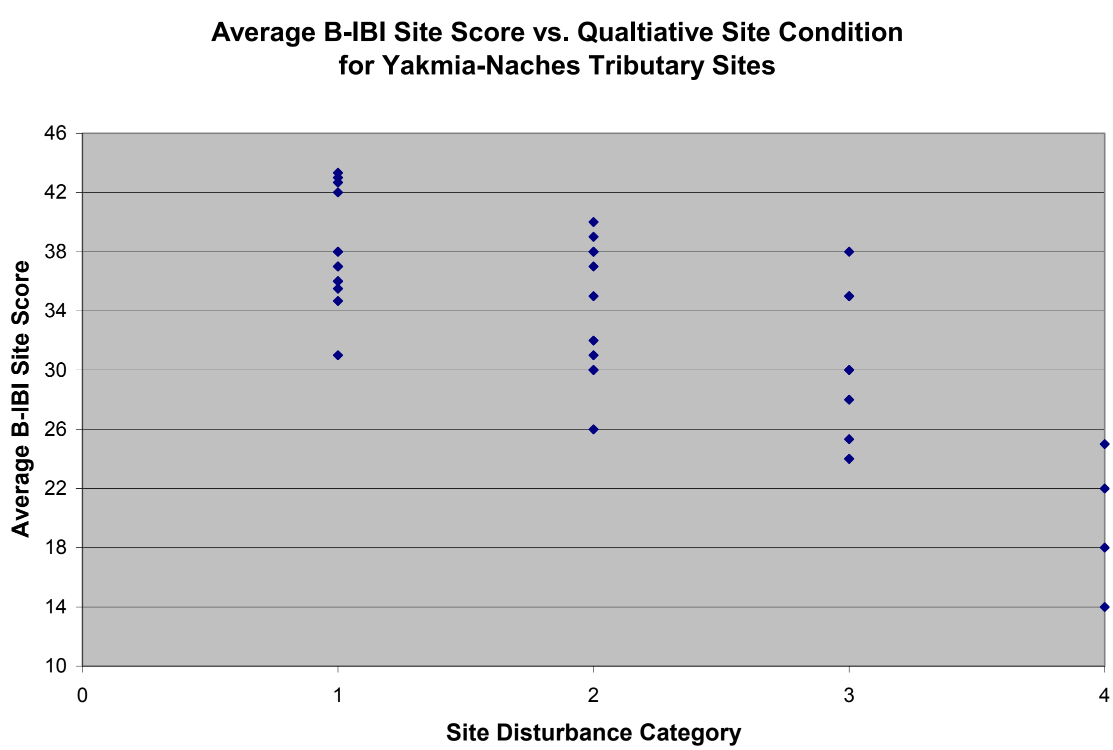
Site-Code	Site Description (Number of years sampled)	Years sampled	Elevation (ft.)	Disturbance Condition	Average B-IBI Score (s.d.)
BIG2	Big Creek, lower (1)	2002	2200	1	38
AHT2	South Fork Ahtanum (3)	2000-2002	2500	1	43 (5.0)
RAT1	Rattlesnake upper (4)	2000-2004	2700	1	36 (1.9)
NAN3	Naneum, middle (1)	2002	2800	1	42
LNA1	M. F. Little Naches (1)	2000	3100	1	36
AME2	American above Hell's Crossing (2)	2001,2002	3300	1	31 (1.4)
COO1	Cooke Creek (3)	2000-2002	3500	1	43 (3.1)
NAN1	Naneum, upper (2)	2000,2002	3500	1	38 (5.7)
BIG1	Big Creek, upper (3)	2000-2002	3600	1	35 (1.2)
NTA1	North Fork Taneum 1 (2)	2000,2002	3700	1	37 (7.1)

NTA2	North Fork Taneum 2 (2)	2001,2002	3700	1	43 (4.2)
MAN1	S.F. Manastash upstream (2)	2000,2002	4300	1	37 (1.4)
AHT1	North Fork Ahtanum (2)	2000,2002	4700	1	36 (2.8)
RAT2	Rattlesnake lower (1)	2000	2000	2	30
TAN2	Taneum @ BOR site (2)	2001,2002	2000	2	39 (4.2)
NIL1	Nile Creek (2)	2000,2002	2400	2	37 (1.4)
OAK1	Oak Creek (2)	2000,2002	2500	2	35 (7.1)
WEN1	N.F. Wenas (2)	2000,2003	2500	2	31 (7.1)
LNA2	Little Naches (1)	2001	2600	2	38
UMT1	Umptanum (1)	2000	2600	2	26
AME3	American at Bumping Xing (1)	2002	2700	2	38
TAN1	Taneum at Taneum C.G. (1)	2000	2700	2	30
MAN2	S. F. Manastash, downstream (2)	2000,2001	2800	2	40 (5.7)
AME1	American at Hell's Crossing C.G. (1)	2000	2900	2	32
AHT3	Lower Ahtanum at Military Museum (3)	2000,01,03	1000	3	25 (4.2)
SBE1	Eschbach Springbrook (1)	2001	1300	3	24
COW1	Cowiche Creek (3)	2001-2003	1500	3	24 (2.0)
TAN3	Taneum below I-90 (2)	2001,2003	1900	3	28 (0)
TEA2	Teanaway Red Rd. Br. (2)	2001,2002	1900	3	35 (4.2)
TAN4	Taneum below BOR (1)	2002	2000	3	30
CAB1	Cabin Creek (3)	2001-2003	2300	3	38 (4.0)
TEA1	N. F. Teanaway (2)	2000,2002	2400	3	35 (9.9)
NAN2	Naneum, Charleton Rd. (1)	2001	2600	3	30
MOX1	Moxee Drain (1)	2000	1000	4	14
WID1	Wide Hollow (2)	2000,2003	1000	4	22 (2.8)
WEN2	Wenas lower (3)	2001-2003	1200	4	18 (2.0)
CHE1	Cherry Creek (2)	2000,2003	1500	4	25 (4.2)

Table 11 Average B-IBI Scores for all tributary site-years for each disturbance category and for all sites and years combined. The number of site-years for the mean scores are given in parentheses.

Type of Site	Mean B-IBI Score (number of site-years)	Site-Score Standard Deviation
Reference	38.0 (28)	4.65
Condition 2	34.9 (16)	5.37
Condition 3	30.1 (18)	6.42
Condition 4	20.3 (8)	4.46
All Sites and Years	33.2 (70)	7.67

Figure 2. Tributary Site Condition and Average B-IBI Site Score. Some points on the chart include multiple sites.



If the scores at each of the three sites had differed by only ten points, their standard deviations would have been 5.7 and the standard deviation of site scores for sites in each of the four condition categories would have been less than six. A standard deviation in B-IBI score of six would permit the construction of 50% confidence intervals (CIs) of plus- or minus-4 points, narrow enough to distinguish five categories of site condition, with break-points between adjacent categories at scores of 18, 26, 34, and 42. We believe that this would provide a reasonable standard for monitoring and assessment purposes (see Discussion).

4.1.3 Interannual Variability in Site Scores

An informative and reliable condition index should display a minimum of interannual variation at sites that remain essentially unchanged between years. It is important, therefore, to examine the variability of site scores in some detail and, in particular, those sites that exhibited the greatest range of scores but at which no obvious disturbances occurred over the period of sampling. We approach this issue by first examining the variability of the component metrics.

Metric scores can take only three distinct values (1, 3, or 5). Consequently, between-years scores can differ from one another only by values of 2 or 4, corresponding to variation by one or two metric steps, respectively. Variation by one step in metric value (1 to 3, 3 to 5, or vice versa) is not surprising. Variation by two steps (1 to 5 or 5 to 1) is expected to be relatively uncommon, if not rare. Of the 37 tributary sites, 24 were sampled in two or more years (Table 10). With 24 sites and ten metrics there are 240 (24 x 10) range values, each 0, 2, or 4. Of the 240 actual metric range values, 21 or 9% displayed the highest range of inter-annual variation in metric score of 4. The remaining 219 had values of 0 or 2. These 21 values were distributed among five metrics, as shown in Table 12.

Table 12. Number of multi-year same-site tributary samples for individual metrics with maximum metric range values of 4 for 24 sites with multi-year samples. Possible range values are 0, 2, or 4. Total number of individual metric range values = 240 (10 metrics times 24 multi-year sites). Numbers in parentheses in Category and Metric Totals headings are the total numbers of multi-year metric range values in each disturbance category and among all 24 sites. Percentages in parentheses in the Category and Metric Totals are the percentages of range values of 4 in the total number of metric range values within each category; for example, among category 1 sites, 9 out of 100 total range values had values of 4.

Metric	Category 1 (100)	Category 2 (50)	Category 3 (60)	Category 4 (30)	Metric Totals (240)
Stonefly Taxa	0	1	0	0	1 (0.4%)
Intolerant Taxa	0	1	2	0	3 (1.3%)
Percent Predators	5	1	3	1	10 (4.2%)
Clinger Taxa	4	1	1	0	6 (2.5%)
Percent Tolerants	0	0	1	0	1 (0.4%)
Category Totals	9 (9%)	4 (8%)	7 (11.7%)	1 (3.3%)	21 (8.8%)

Of the 21 interannual metric scores with range-values of 4, 19 (8% of the 240 total range values) were distributed among three metrics: number of intolerant taxa, total number of clinger taxa, and percent predator individuals. The preponderance of the variation is in the percent predator and clinger taxa metrics. Together the variation in these two metrics represent a total maximum metric range variation of less than 7% (16 of 240) of the total range variation possible.

At the sample sizes involved there are no significant differences between the proportions within disturbance categories or between the proportions within a category and the mean proportion over all four categories (n = 240; exact binomial probabilities test, not shown). In other words, these metrics appear to be behaving reasonably well.

The variability of the three most variable metrics (Intolerant Taxa, Percent Predators, and Clinger Taxa) is more appropriately assessed in the context of all ten metrics that together constitute the Index score, because the value of a multi-metric index lies in its ability to reflect several different kinds of relevant biological signal at a site and integrate them into a single measure. Informative metrics that display relatively high degrees of variability should be balanced by the lower variability of the remaining metrics. Consequently, the more important variability to examine is the variability of site scores themselves.

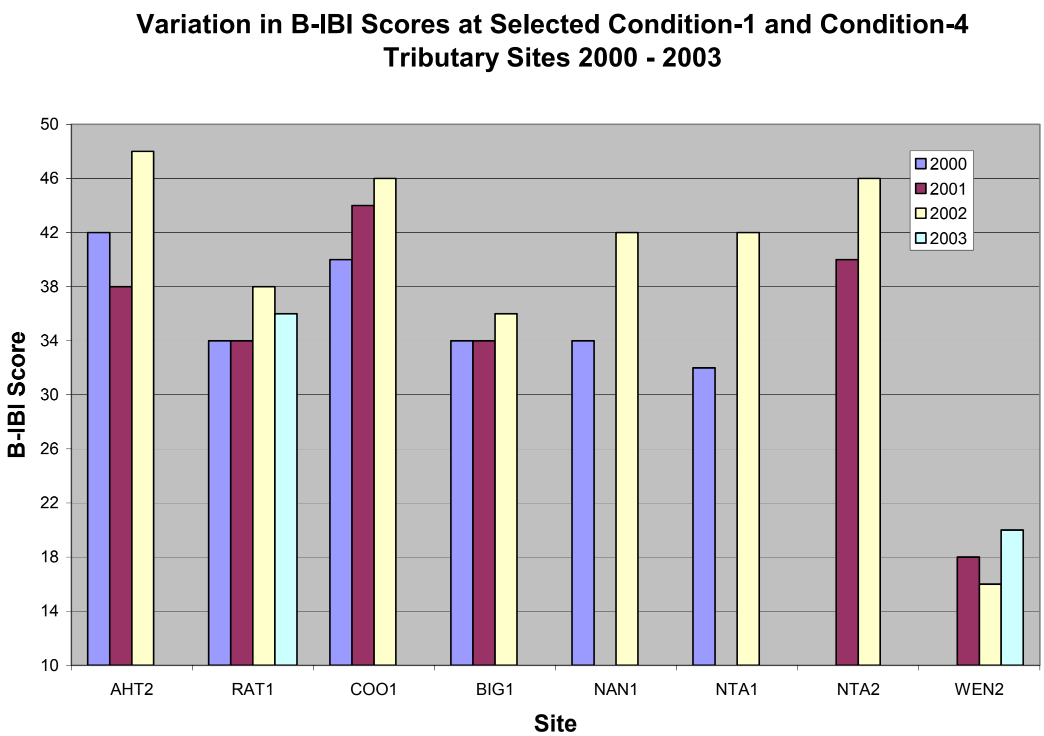
4.1.4 Patterns of Variation of Selected Tributary Sites

Figures 3 and 4 show the B-IBI site scores for 18 of the 24 tributary sites with multiple years of sampling spanning all four qualitative site conditions. Figures 5 – 7 show the metric scores for the three metrics with the most frequent range-4 values for all tributary sites with multiple years of sampling whose standard deviations in B-IBI scores was greater than 3.0 (n = 14; Table 10).

Site AHT2, the South Fork of Ahtanum Creek, is a reference site. It was sampled near South Ahtanum Road approximately five miles above Tampico in 2000, 2001, and 2002. IBI score varied from 38 to 48, with a standard deviation of 5.0 (Table 10, Figure

Figure 3 and 4. Variation in B-IBI scores of selected condition-1 and condition-4 tributary sites and selected condition-2 and condition-4 tributary sites.

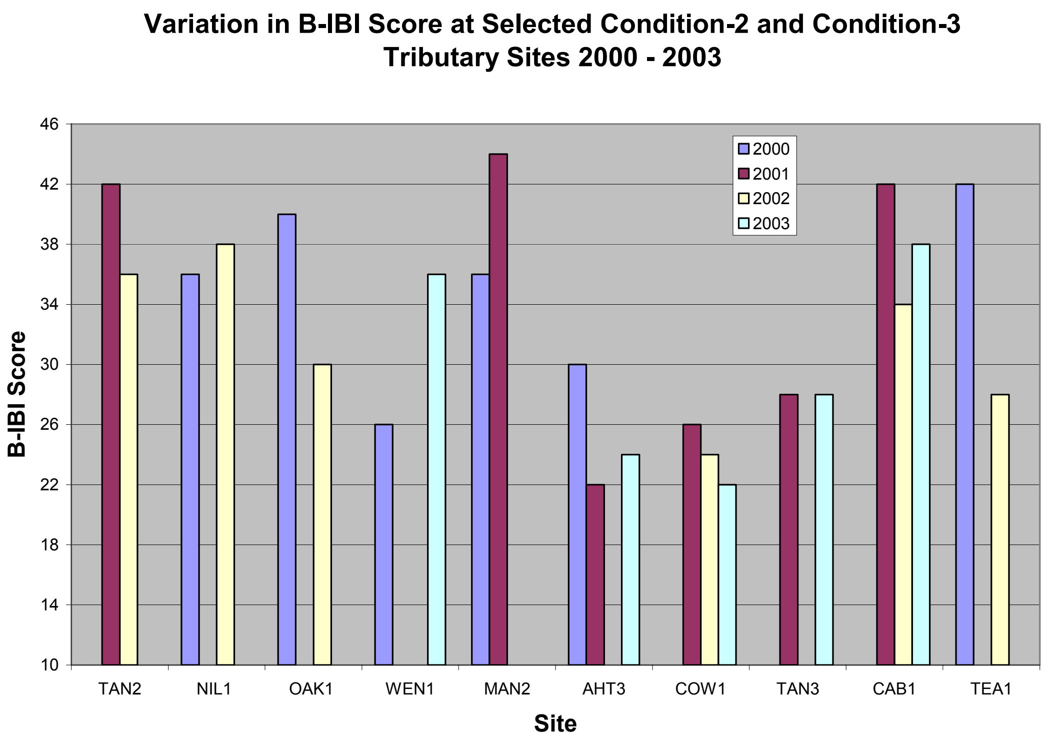
Figure 3.



3.). The score of 48 in 2002 was the single highest site score of the entire data set. Two of the three most-variable metrics showed the maximal range of variation of 4, Clinger Taxa (Figure 5) and Percent Predators (Figure 7). The third metric ranged between values of 3 and 5. Only one of the three (Percent Predators) scored the lowest value of 1 in the year in which the site IBI score was lowest (2001) and neither had the lowest value in the same year. Consequently, small ranges of variation (range = 2) in several metrics are responsible for the observed variation in IBI score at this site, which is as it should be.

COO1, upper Cooke Creek, is on Washington State Department of Natural Resources (WDNR) land. It had a relatively low standard deviation of site score (3.1, Table 10) and scored consistently high over three years of sampling. It, too, displayed the maximum range of variation for both the Clinger Taxa and Percent Predator metrics and showed a similar temporal pattern to AHT2. Neither of the metrics scored a 1 in the same year. Again, the variation was largely due to small ranges of variation in several metrics.

Figure 4.



We established two reference-quality sites within 250 - 300 meters of one another on upper North Fork Taneum Creek, NFT1 and NFT2. Each was sampled in two years (2000,2001 and 2001,2002, respectively).The main stream channel at both sites is densely loaded with large woody debris of several species of old growth evergreens, and have heavily wooded, mature riparian areas.

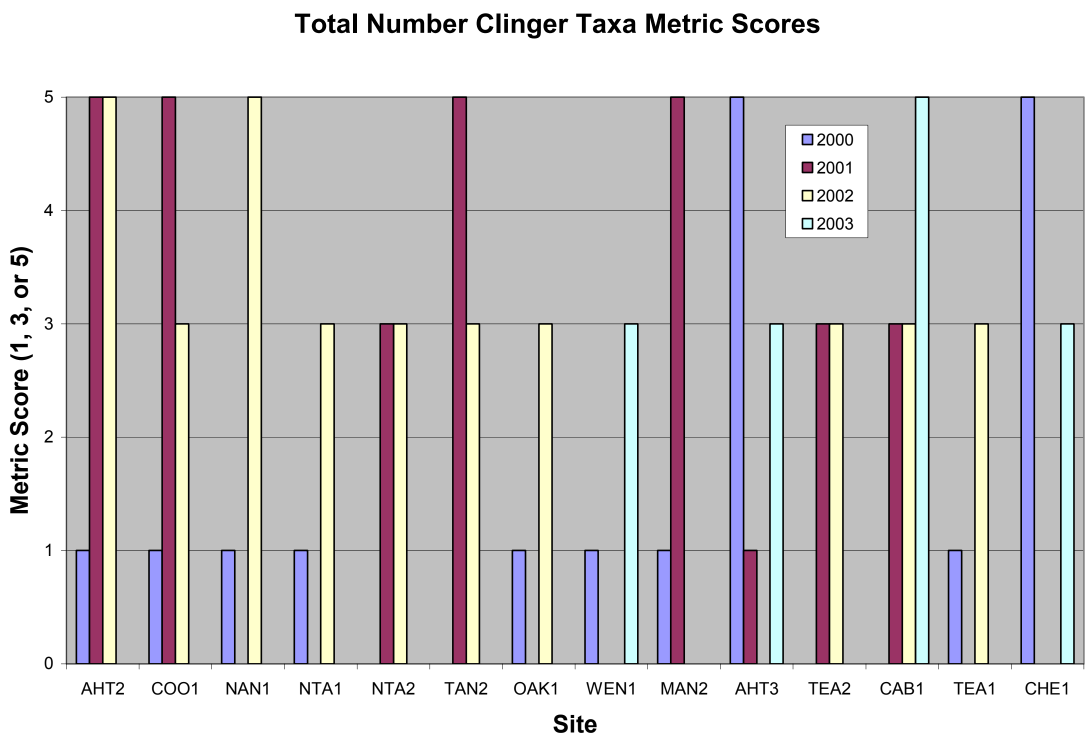
NFT1 was located over 200 meters downstream of a well-designed and well-built road crossing culvert. So we expected the influence of the road crossing to be minimal. The site scored lower than expected (32) in the first sample year (2000). Both the Percent Predator and the Intolerant Taxa metric, however, scored 5, while the Clinger Taxa metric scored 1. In addition to the Clinger Taxa metric the Dominance metric also scored 1 in this year, due in part to a surprisingly large number of oligochaetes in the sample. This suggested to us that perhaps fine sediment from the logging road was having a detectable impact on the site, so we established a second site, NFT2, over 50 meters upstream of the road crossing.

NFT 2 scored in the expected range (40 and 46) in both years as did NFT 1 (42) in the second year of sampling (2001). Among the three most-variable metrics, only NFT2 showed a range of variation of 4 and this only for Percent Predators, the single most-variable metric overall (Figure 7). It is possible that by chance we encountered the legacy of a pulse of fine-sediment input at NFT1 during the first year of sampling, in which case the B-IBI detected this departure from reference-condition. Replication nearby in space and in time confirmed the initial characterization of the stream segment as a reference stream. In addition, high scores for Intolerant Taxa and Percent Predators metrics highlighted the potential high biological integrity of the site.

Figure 4 shows B-IBI scores for five sites classified as condition-2 (TAN2, NIL1, OAK1, WEN1, and MAN2) and five sites classified as condition-3 (AHT3, COW1, TAN3, CAB1, AND TEA1). Four of the five condition-2 sites (TAN2, OAK1, WEN1, and

Figure 5 – 7. Between-Year Variation Metric Scores for the Three Most-Variable Metrics for all Tributary Sites with Two or More Years of Sampling and Standard Deviations in B-IBI Score Greater than 3.0. See Table 8 for site-name abbreviations.

Figure 5.

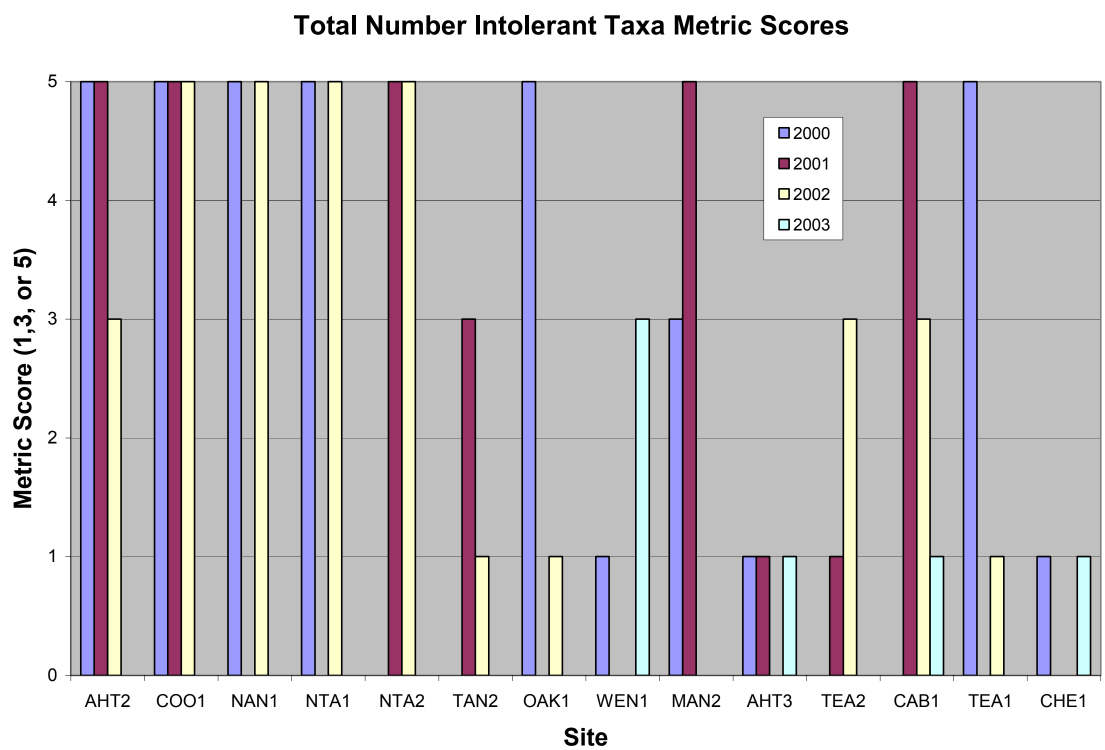


MAN2) had standard deviations in B-IBI score greater than 3.0. Two of these four, MAN2 (South Fork Manastash Creek upstream of the Manastash Creek Road crossing) and OAK1 (Oak Creek within the Washington Department of Fish And Wildlife Oak Creek Wildlife Area) had a range of variation of 4 in one or more of the three most-variable metrics. For MAN2 this occurred in both the Clinger Taxa and Percent Predator metrics (Figures 5 and 7). As in the case of the condition-1 streams discussed above, in no year did the site score the lowest value (1) in both metrics. OAK1 (Oak Creek within

the Washington Department of Fish And Wildlife Oak Creek Wildlife Area) had the maximum range of metric score variation only for the Intolerant Taxa metric (Figure 6).

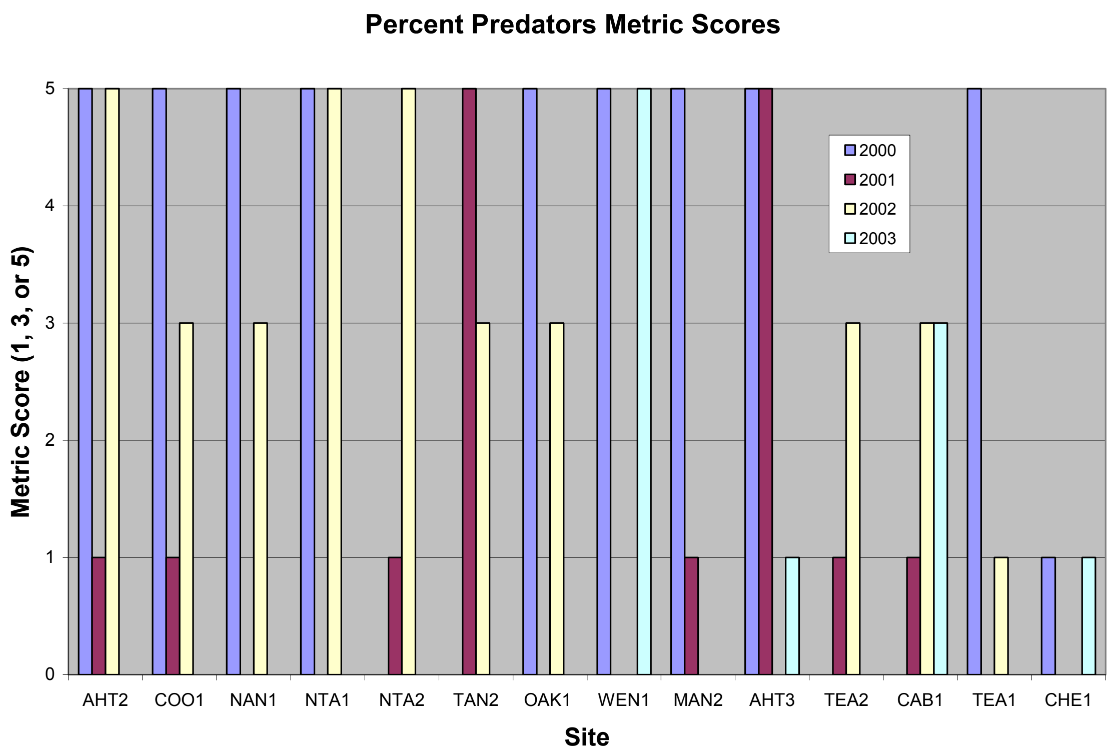
Three of the five condition-3 sites shown in Figure 4 had standard deviations in B-IBI scores greater than 3.0 (AHT3, CAB1, and TEA1). AHT3 and TEA1 showed a maximum range of variation in metric score for two of the three most-highly variable metrics.

Figure 6.



AHT3 (Ahtanum Creek on the Central Washington Agricultural Museum grounds) showed a maximum range for Clinger Taxa and Percent Predators (Figures 5 and 7). TEA1 (North Fork Teanaway river above the North Fork Road bridge) showed a maximum range for Intolerant Taxa and Percent Predators (Figures 6 and 7). CAB1 (Cabin Creek above Easton, downstream of Cale Creek and the washed-out Forest Road

Figure 7.



41 bridge) showed the full range of variation for the Intolerant Taxa metric (Figure 6).

Only TEA1 broke the pattern of only scoring the lowest value of 1 for a single high-variation metric in any one year. The site scored 1 for both Intolerant Taxa and Percent Predator metrics in 2002 and scored a 5 for both in 2000. The year with low scores for both of these metrics was also the year that the site scored lowest in IBI (28), and the year with high scores was also the year with the highest IBI score of 42, an anomalously high score for a condition-3 site. The anomalous character of this score (as well as the score of 42 in 2001 at the CAB1 site) is discussed in the next subsection on the correspondence of B-IBI score to site classification. But with the exception of this site, we conclude that the between-year variation in B-IBI score at a site is not unreasonably large and, consequently does not indicate unreliable performance of the multi-metric index (Table 9).

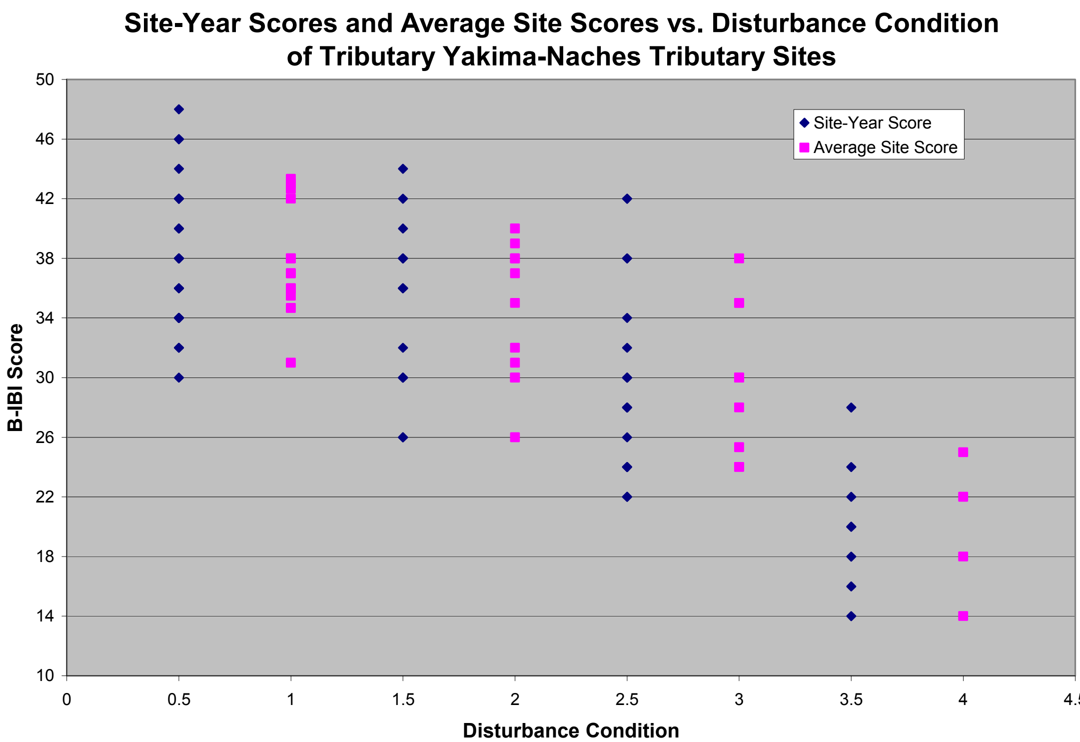
4.1.5 Correspondence Between Qualitative Site Classification and B-IBI Score

One of the primary purposes of developing an assessment index such as the B-IBI is to help refine our biological intuition regarding stream condition at the reach and larger spatial scales. We used our biological intuition, based upon our knowledge of stream ecology, to develop qualitative/semi-quantitative narrative criteria that we employed objectively to the best of our abilities to characterize the condition of our sampling sites on an ordinal scale. We then used the initial classification of sites to evaluate metrics and refine metric scoring criteria. Finally, as discussed in the preceding section we evaluated the performance of the metrics and the aggregate index by examining the variability of site scores, under the assumption that within the period of sampling (2000 to 2003) conditions at each site did not change significantly.

The final matter to examine is how well the yearly and average multi-year IBI scores at sites corresponds to the initial ordinal-scale characterization of the sites. We expect some discrepancy and are willing to use the scores to correct some of our initial qualitative classifications, provided the discrepancies are neither too great nor too many. This requires that we first examine condition-1 and condition-2 sites that score unexpectedly low in one or more years, and second that we examine condition-2 and condition-3 sites that score unexpectedly high. If we can reconcile all or most such discrepancies, we can present a revised and, tentatively, final list of sites and site-conditions.

Figure 8 shows the relationship between tributary site condition and average B-IBI site score, and site condition and site-year score. To show both on a Microsoft Excel chart numbers designating site conditions for site-year scores have been recoded by subtracting 0.5 from each category, so that 0.5 is equivalent to 1, 1.5 is equivalent to 2, and so forth.

Figure 8. Qualitative tributary site condition vs. (a) average site B-IBI scores and (b) site-year B-IBI scores. Numbers designating conditions for site-year scores have been recoded by subtracting 0.5 from the original numeric codings to enable Microsoft Excel's scatter plot chart function to display both sets of data.



The highest tributary B-IBI score recorded was 48 and the lowest was 14, for a range of 34. Dividing this range into quartiles would yield a quartile range of 8.5. If we had equal numbers of sites and site-years in each of the qualitative conditions, we might reasonably expect the B-IBI scores to fall into quartiles each with a range of 8.5. Rounding to 9 and starting with condition 1, this would yield expected range for each site condition of 39 – 48, 30 – 39, 21-30, and 14 – 21.

It is generally more difficult for a site to score high than low. In addition the data set has more sites in conditions 1 and 2 (13 and 11, respectively) than in conditions 3 and 4 (9 and 4, respectively). So it makes sense to be more lenient in the expected range of condition 1 and 2 sites than for conditions 3 and 4. This also errs in the direction of leniency by making it more likely that a site will get classified as a 1 or 2 than a 3 or 4. So we might expect condition 1 site to lie within the range of 38 to 48, condition 2 sites within the range 28 to 37, condition 3 sites within the range 20 to 27 and condition 4 sites to be below 20. Using these adjusted ranges as rough guidelines we should puzzle about condition 1 sites that score lower than 38, condition 2 sites that score below 28 or above 37, condition 3 sites that score below 20 or above 27, and condition 4 sites that score above 20.

Table 13 lists these revised scoring criteria for site-classification.

Table 13. Recommended Score-Based Site Condition Classification Criteria

B-IBI Site Score Range	Score-based Site Condition Classification
38 – 50	1
28 - 37	2
20 - 27	3
10 - 19	4

Using the adjusted scoring criteria, of a total of 13 condition-1 sites (by narrative criteria) and 28 site-years, there are eight sites accounting for 13 site-years that have scores below 38 (range 30 to 36; Appendix Table A1, Appendix Figures A1 - A4). Ten of the 13 sites have multiple years of sampling and of these ten, seven have average scores below 38 (range 31 to 37; Table 14).

By narrative criteria, there are 11 condition-2 sites accounting for 16 site-years. Of these, five sites (six site-years) are classified as condition-1 by the scoring criteria and two sites accounting for two site-years are classified as condition-2 by the scoring criteria (Appendix Table A2, Figures A1 – A4). There are five sites with multiple years of sampling. One of these is classified as condition-1 and one is classified as condition-3 (Table 14).

By narrative criteria, there are nine condition-3 sites accounting for 18 site-years. Of these, none were classified by the scoring criteria as being in worse condition. Seven sites accounting for 12 site-years were classified as being in better condition than the narrative classification. Eight of the 12 site-years were classified as condition-2 and the remaining four as condition-1 (Table A3, Figures A1 = A4). There are six sites with multiple years of sampling, one of which is classified as condition-1 by the scoring criteria and three of which are classified as condition-2 (Table 14).

By narrative criteria, there are five conditon-4 sites accounting for eight site-years. Of these, four site-years were classified as condition-3 by the scoring criteria and one was classified as condition-2. There are three sites with multiple years of sampling, two of which are classified as condition-3 by the scoring criteria (Table 14).

Over all 37 tributary sites and 70 site-years, 32 site-years (23 sites) showed agreement (difference = 0) between the initial qualitative, narrative classification of site condition and the final recommended score-based classification. Among the 13 sites and 28 site-years classified as condition-1 by narrative criteria, ten sites (15 site-years) showed agreement (difference = 0) with the score-based classification. Among the eleven sites and 16 site-years classified as condition-2 by narrative criteria, eight sites (eight site-years) showed agreement. Among the nine sites and 18 site-years classified as condition-3 by narrative criteria, three sites (six site-years) showed agreement. Among the five sites and eight site-years classified as condition-4 by narrative criteria, two sites (three site-years) showed agreement.

Twenty-four tributary sites were sampled in two or more years. Data for these sites is summarized in Table 14. Table 14 lists the classifications of the 24 tributary sites with multiple years of sampling ordered on the basis of average site scores. Table 15 lists all 37 tributary sites ordered on the basis of average site score and Figure 9 plots average site scores and condition class against site elevation.

Table 14 Comparison of narrative and mean score-based site condition for 24 tributary sites with multiple sample years (n = 2 – 4) A negative difference means that the scoring criteria classify the site as being in poorer condition than the narrative classification. A positive difference means the scoring criteria classify the site as being in better condition than the narrative classification.

Site-Code	Mean B-IBI Site Score	Narrative (N) Condition	Score-Based (S) Condition	Difference (N - S)	Elevation
AHT1	36	1	2	-1	4700
AHT2	43	1	1	0	2500
AME2	31	1	2	-1	3300
BIG1	35	1	2	-1	3600
COO1	43	1	1	0	3500
MAN1	37	1	2	-1	4300
NAN1	38	1	1	0	3500
NTA1	37	1	2	-1	3700
NTA2	43	1	1	0	3700
RAT1	36	1	2	-1	2700
MAN2	40	2	1	1	2800
NIL1	37	2	2	0	2400
OAK1	35	2	2	0	2500
TAN2	39	2	1	1	2000
WEN1	31	2	2	0	2500
AHT3	25	3	3	0	1000
CAB1	38	3	1	2	2300
COW1	24	3	3	0	1500

TAN3	**28**	**3**	**2**	**1**	**1900**
TEA1	**35**	**3**	**2**	**1**	**2400**
TEA2	**35**	**3**	**2**	**1**	**1900**
CHE1	**25**	**4**	**3**	**1**	**1500**
WEN2	**18**	**4**	**4**	**0**	**1200**
WID1	**22**	**4**	**3**	**1**	**1000**

Table 15. Score-based classification of all 37 tributary sites. Sites with multiple years of sampling in bold font.

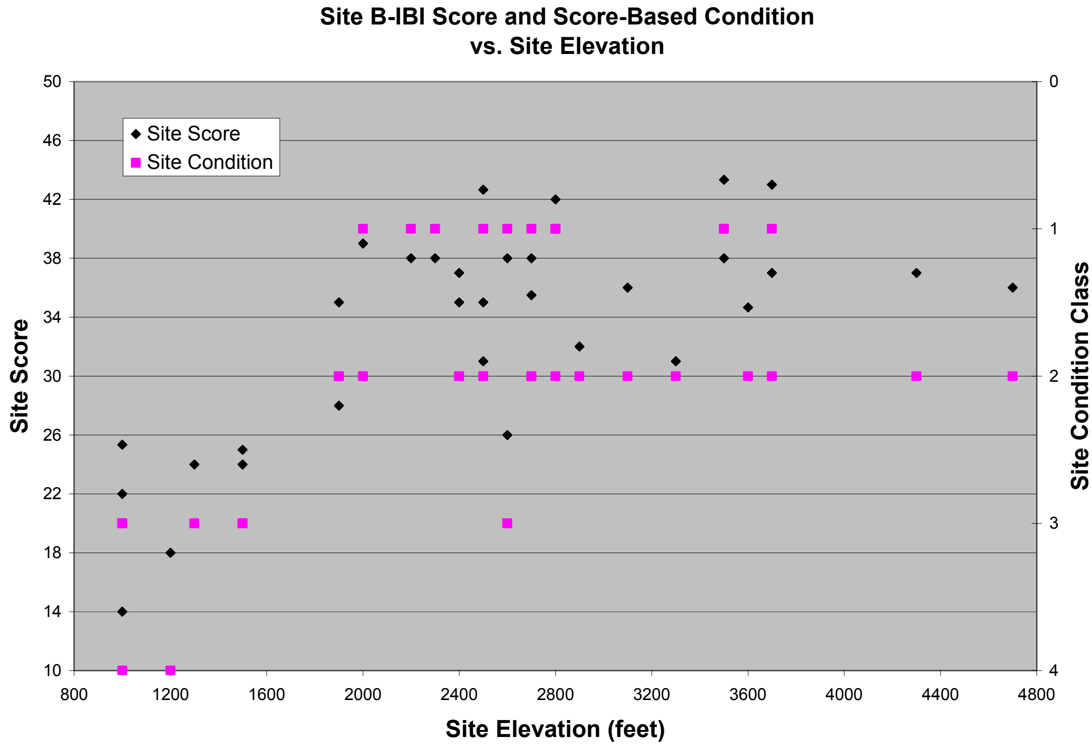
Site-Code	Score-Based (S) Condition	Site Elevation (feet above mean sea level)
AHT2	**1**	**2500**
AME3	**1**	**2700**
BIG2	**1**	**2200**
CAB1	**1**	**2300**
COO1	**1**	**3500**
LNA2	**1**	**2600**
MAN2	**1**	**2800**
NAN1	**1**	**3500**
NAN3	**1**	**2800**
NTA2	**1**	**3700**
TAN2	**1**	**2000**
AHT1	**2**	**4700**
AME1	**2**	**2900**
AME2	**2**	**3300**
BIG1	**2**	**3600**

LNA1	**2**	**3100**
MAN1	**2**	**4300**
NAN2	**2**	**2800**
NIL1	**2**	**2400**
NTA1	**2**	**3700**
OAK1	**2**	**2500**
RAT1	**2**	**2700**
RAT2	**2**	**2000**
TAN1	**2**	**2700**
TAN3	**2**	**1900**
TAN4	**2**	**2000**
TEA1	**2**	**2400**
TEA2	**2**	**1900**
WEN1	**2**	**2500**
AHT3	**3**	**1000**
CHE1	**3**	**1500**
COW1	**3**	**1500**
SBE1	**3**	**1300**
UMT1	**3**	**2600**
WID1	**3**	**1000**
MOX1	**4**	**1000**
WEN2	**4**	**1200**

Table 14 reveals that ten of the 24 sites with multiple years of sampling show agreement (difference = 0) between the initial narrative classification of site condition and the score-based classification. Six had a difference of -1 (narrative condition higher than score-based), seven had a difference of +1 (score-based condition higher than narrative), and one had a difference of +2. Five of the six sites with a difference of minus-1 were at elevations greater than 3200 feet. All involve sites classified as condition-1 (reference quality) sites by narrative criteria but classified as condtion-2 by IBI score. These cases likely indicate that the scoring criteria exaggerate the extent to which these sites depart from a reference condition.

The seven sites with a difference of plus-1 are located at elevations at or below 2800 feet, and include upgrades of sites classified by narrative criteria as condition-4, condition-3, and condition-2. There are two possibilities to explain these discrepancies. Our scoring criteria may be too lenient and thereby fail to reflect the full extent of impairment at these sites; or, our narrative criteria may be too strict or otherwise fail to incorporate features of sites that reflect elements of biotic integrity. In light of the discrepancy in the opposite direction for the six sites classified as condition-1 by narrative criteria, we believe that the first alternative (narrative criteria too strict) is less likely than the second. That is, we think it reasonable to believe that the scoring criteria are by-and-large accurate and indicate the biotic integrity of site condition more accurately than our narrative criteria. Accordingly, we propose that the list of site conditions given in Table 15 be tentatively accepted as accurate.

Figure 9. Tributary site elevation (feet above mean sea level) and average B-IBI site scores and Disturbance Categories.



There remains one discrepancy that requires additional discussion. The site on lower Cabin Creek in the upper Yakima subbasin received a narrative classification of condition-3, but was classified by B-IBI score as condition-1 on the basis of a three-year average score of 38. This site experienced a major flood event in the winter of 1996/97 that scoured the stream channel and the associated riparian zone, and resulted in a shift of the main stream channel from one side of the channel migration zone to the other. When first sampled in 2001 the riparian zone was essentially void of shrubs and trees with the exception of scattered seedling cottonwood and alder and the stream channel was fully exposed and unshaded. The channel itself had modest sinuosity, gravel/cobble substrate, and minimal fines and embeddedness, and was dominated by shallow riffles with few pools. Median grain size based on a Wolman pebble count was 58 mm and water depth averaged less than 30 cm. The lack of channel depth, pools, and canopy cover indicate poor fish habitat which supported a narrative classification of condition-3.

As shown in Table 14, this site scored as condition-1 in 2001 and 2003, and as condition-2 in 2002. The surprising high score of 42 on the initial sampling occasion in 2001 prompted sampling in each of the following two years. Although this additional sampling resulted in some variability (range 34 to 42, Table 14), the site clearly revealed more biological integrity than met the eye. We are inclined to accept the results as telling us that the site has significant biological potential. Other features of stream habitat are limiting for salmonids, but it does not appear that the food web is.

4.2 Response of Tributary Sites to Disturbance Factors

Having evaluated metrics and scoring criteria and determined the classification of site condition on the basis of site score (Table 19), we examined selected landuse data to evaluate the response of tributary sites to types of landuse. Landuse data examined included GAP data on road densities, percent forest cover, percent agricultural land cover, and percent developed land cover. The coverage for developed land cover at the subbasin scale lacked the detail necessary for detecting any clear signal using B-IBI. Percent developed land cover ranged from 0 to 1 % in most tributary subbasins to a high of 4% in the Wide-Hollow Creek subbasin. A more detailed analysis of developed land cover was not pursued, so we report no results for this category.

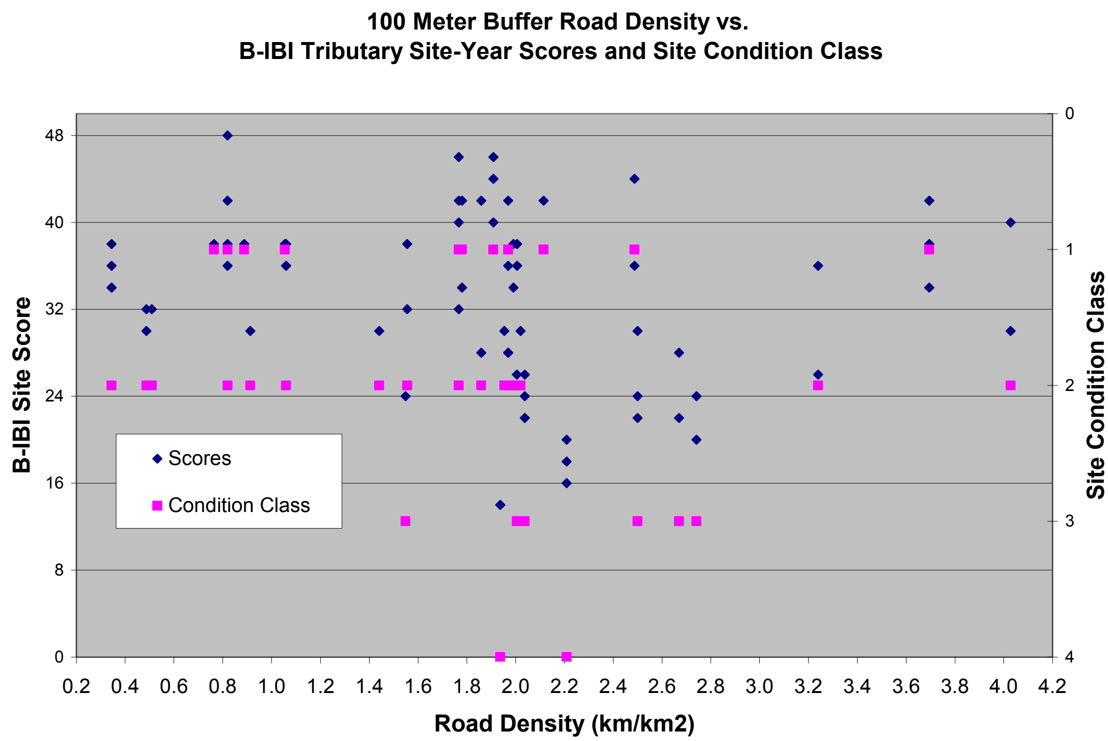
4.2.1 Road Density and Tributary Site Condition

In Figure 14 site-year scores and site condition class are plotted against road density within a buffer of 100 meters on each side of the each tributary stream. Road density is reported as length of road in kilometers per square kilometer of catchment area. Catchment area is measured for the area within the 100 meter buffers upstream of the sampling site based upon Washington DNR GAP data for 2000**.** Data for upper Big Creek (BIG1) was not available. For the remaining 36 tributary sites, road densities ranged from a low of 0.34 for upper Rattlesnake Creek in the Naches subbasin to 4.03 for Oak Creek in the Naches/Tieton subbasin, with a mode (7 of 36) at 2.0. For 19 of the 36 sites (53%) road densities were between 1.6 and 2.4 k/k^2.

Figure 14 shows that road density displays no clear correlation with either site-year score or site condition class. Within the modal range of 1.6 to 2.4, all site conditions occurred and site-year scores were widely distributed, covering a range from the lowest score for any site-year (14 at Moxee Drain) to 46.

The only potentially significant feature of the data is that no site with a road density of 1.2 or lower had a site-year score below 30 or a site classification lower than 2. This accounted for 15 of 67 site-year scores and condition class at nine of 36 sites. Clearly, road density *per se* does not account for these site scores. Road density interacts with other landscape features and conditions that we were not able to analyze, In addition, the road density data does not identify road conditions, and we were thus not able to evaluate the impact of road condition on site scoring.

Figure 14. Tributary Site Classification by B-IBI Score vs. Road Density within 100 Meter Stream Buffers For 36 Tributary Sites and 67 Site-Years.



In general, it does appear that at the maximum 100-meter buffer densities observed in the GAP data that we had access to, landform and riparian condition appear to protect tributary streams throughout most of the upper Yakima and Naches subbasins from pronounced impacts typically attributable to roading (excessive fines, including dust, and increased flashiness in the hydrograph due to run-off).

<u>*4.2.2 Percent Forest Cover and Tributary Site Condition*</u>

Land cover was obtained from a Washington Department of Fish and Wildlife Gap Analysis Program (GAP) data set (available online at http://www.wdfw.wa.gov/wlm/gap/dataprod.htm). For the GAP project, land cover polygons were derived from a 1991 Landsat TM image, with each mapping unit (i.e.

polygon) a minimum size of one hectare (0.01 square kilometers). We overlaid the GAP land cover data with our watershed polygons and, using the primary land cover[2] attribute (PRIM), calculated the percent cover for percent forest land cover, percent agricultural land cover and percent developed land cover within each watershed.

Table 16 summarizes the correlations between percent forest cover from the GAP data and site B-IBI score and the individual metrics for all tributary sites. Figure 15 shows the plot of the linear regression of B-IBI site score against percent forest cover for tributary

Table 16. B-IBI and metrics in relation to watershed forest cover for mainstem sites. Shown are R^2 values obtained from linear regression. * = p<0.05, ** = p<=0.01, *** = p<0.001, **** = p<0.0001, ns = not significant.

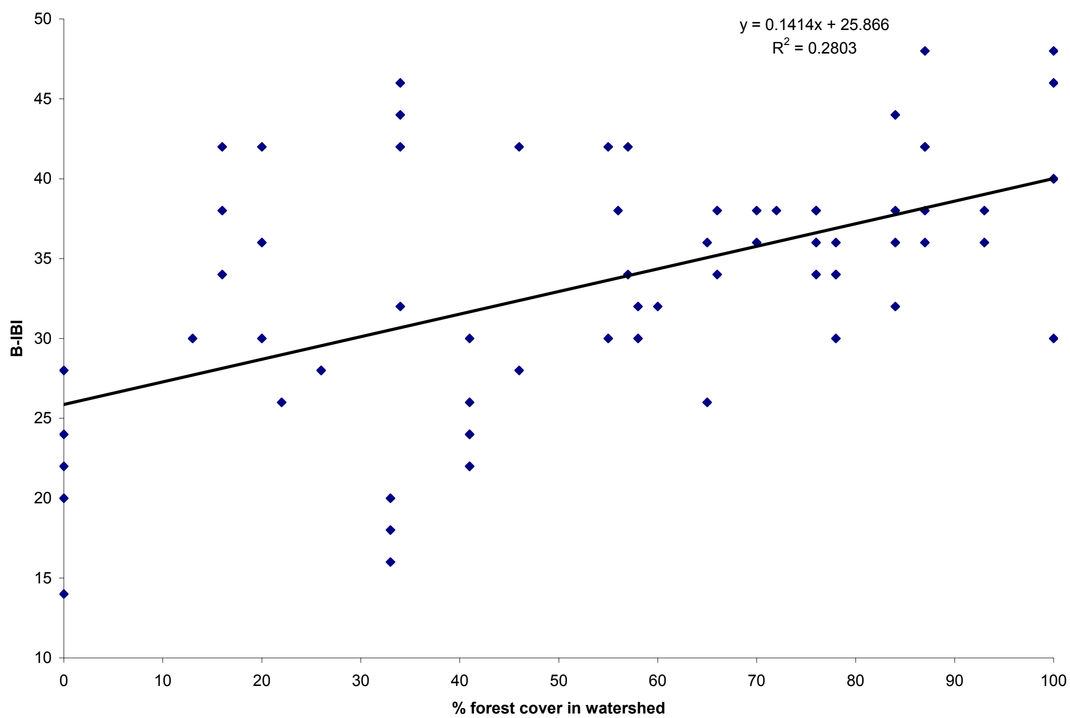
Index	watershed forest cover (R^2)
B-IBI Score	0.28****
total taxa	0.19***
ephemeroptera taxa	0.25****
plecoptera taxa	0.15**
trichoptera taxa	0.21****
long-lived taxa	0.17***
intolerant taxa	0.19***
percent predators	0.015^{ns}
clinger taxa	0.01^{ns}
percent tolerant	0.02^{ns}
percent dominant	0.06*

site-years for which cover data was available. The linear regression of site-year score on percent forest cover was highly significant at an α value of 0.05 ($p<0.001$), with a positive slope of 0.141 and an r-squared value of 0.28. This shows that B-IBI in tributaries responds strongly to percent forest cover: approximately 28 percent of the variance in site score may be accounted for by percent forest cover.

---

[2] Primary land cover in the GAP data set is the type of land cover making up the highest proportion of the total area of each polygon.

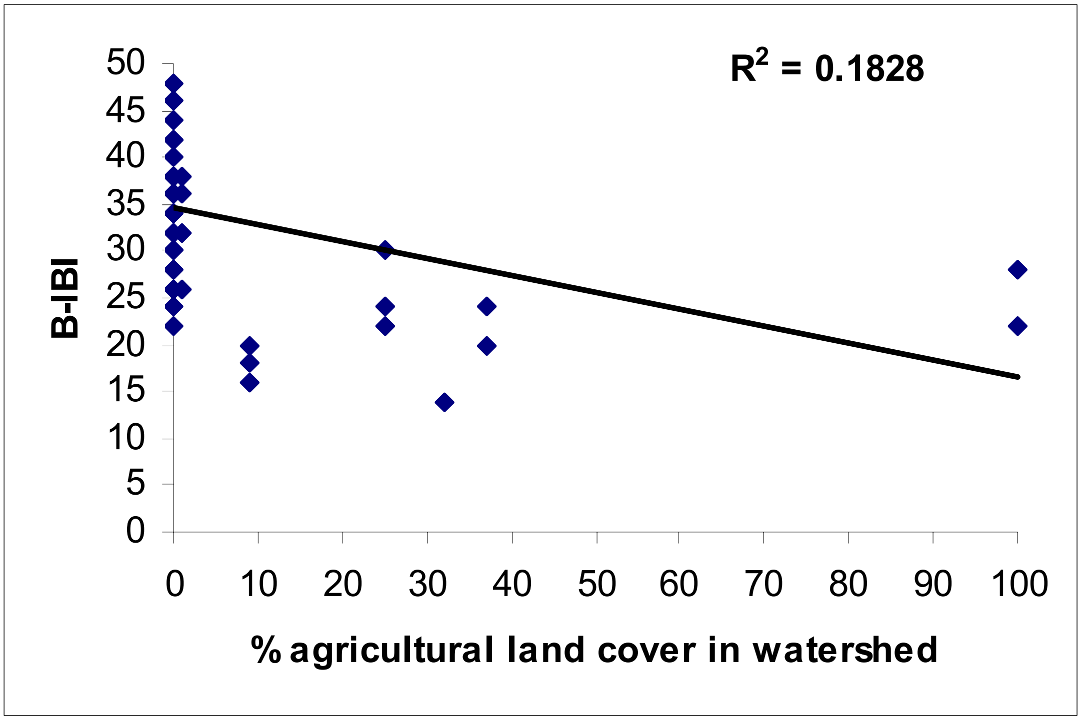
Figure 15. Relationship between B-IBI and watershed forest cover in tributary sites. Slope = 0.141, R2 = 0.28.



<u>*4.2.3 Percent Agricultural Land Cover and Tributary Site Condition*</u>

Figure 16 shows the results of the linear regression of percent of agricultural land cover in tributary subbasins and tributary site B-IBI score. The regression is significant ($p < 0.0001$) with a negative slope of -0.1813, and an R^2 value of 0.1828.

Figure 16 . Percent agricultural land cover and tributary site B-IBI score.



4.3 Mainstem Sites

All mainstem sites, elevations, average B-IBI scores and condition classification are listed in Table 17. Site-year scores and condition classifications are listed in Table 18.

Table 17. Mainstem Sites, Elevations, Average B-IBI Site Scores and Site Condition Class.

Site Code	Site Location Description	Elevation	B-IBI Score	Condition
BUM1	Bumping at Cedar Springs C.G.	2800	43	1
BUM2	Bumping above Bumping Reservoir	3700	34	2
BUM3	Bumping below Bumping Crossing Br.	3400	34	2
CLE1	Cle Elum at Salmon la Sac	2400	22	3
CLE2	Cle Elum above Lake	2300	22	3
CLE3	Cle Elum at Roslyn	2000	22	3
NAC1	Naches at Cottonwood CG	2200	35	2
NAC2	Naches at Naches	1400	33	2
NAC3	Naches at Horseshoe Bend	1700	28	2
NAC4	Naches at Wapatox	1600	30	2
TIE1	Tieton	1900	26	3
YAK01	Yakima at Golf Course Rd.	2100	33	2
YAK02	Yakima @ RRR	1700	25	3
YAK03	Yakima @ Ringer Rd	1400	26	3
YAK04	Yakima @ Big Horn	1400	34	2
YAK05	Yakima @ Moxee Drain	1000	26	3
YAK06	Yakima @ Parker	900	28	2

YAK07	Yakima @ WDFW boat launch	2100	34	2
YAK08	Yakima @ Crystal Springs	2400	30	2
YAK09	Yakima upstream Hwy 24	1000	29	2
YAK10	Yakima downstream Hwy 24	1000	28	2
YAK11	Yakima Upstream Wilson Creek	1400	34	2
YAK12	Yakima downstream Wilson Creek	1400	26	3

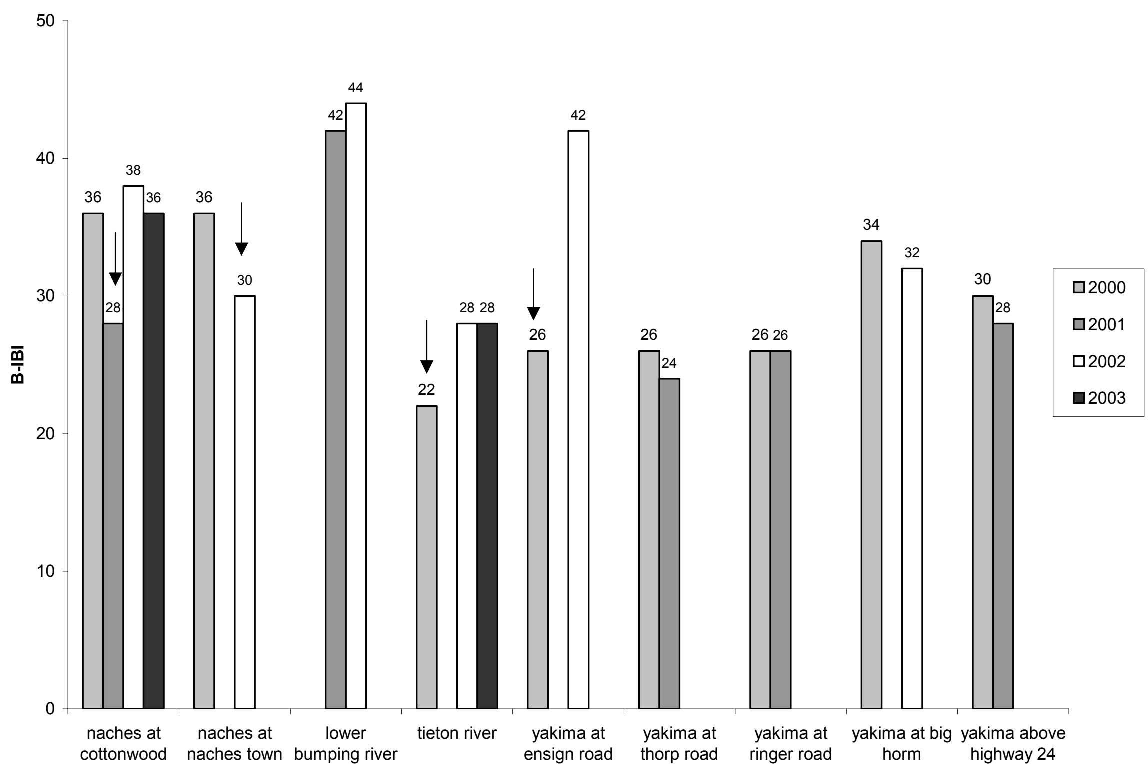
Table 18. Mainstem Sites-Years, B-IBI Site-Year Scores and Site Condition Class Site-Codes in bold font indicate site years[3]

Site-Code and Site-Year	B-IBI Score	Condition
BUM1.01	42	1
BUM1.02	44	1
BUM2.02	34	2
BUM3.02	34	2
CLE1.00	22	3
CLE2.01	22	3
CLE3.01	22	3
NAC1.00	36	2
NAC1.01	28	2
NAC1.02	38	1
NAC1.03	36	2
NAC2.00	36	2
NAC2.02	30	2
NAC3.01	28	2
NAC4.01	30	2
TIE1.00	22	3
TIE1.02	28	2
TIE1.03	28	2
YAK01.00	28	2
YAK02.00	26	3
YAK02.01	24	3
YAK03.00	26	3
YAK03.01	26	3
YAK04.00	34	2
YAK04.02	32	2
YAK05.00	26	3
YAK06.00	28	2
YAK07.01	26	3
YAK07.03	42	1
YAK08.02	30	2
YAK09.01	30	2
YAK09.02	28	2
YAK10.01	28	2
YAK11.02	34	2
YAK12.02	26	3

[3] Site-codes in bold font indicate sites classified in different condition classes in different years in which at least one site-year score differed by more than 4.0 from a classification threshold score (see Discussion and Table 19 below).

Table 18 shows that mainstem site-year scores displayed little inter-annual variation in score and associated condition classification. Only two sites accounting for five of 35 site-years were classified in more than one condition in different site-years and differed in site-year scores by greater than 4.0. The site-year scores at these two sites are shown together with those of seven other sites in Figure 17.

Figure 17. Variation in B-IBI between sample years. Values are displayed for nine mainstem sites sampled across multiple years. Years in which B-IBI differed by more than 4.0 between years at a site are denoted with arrows.



The Tieton River site scored a 22 (condition 3), which was 6.0 below the threshold separating condition 3 and condition 2. The YAK07 site (WDFW Game Access, Ensign

Road at Easton) scored 26 (condition 3) in 2001 and 42 (condition 1) in 2003. These scores at YAK07 were the most anomalous of the entire 105 site-year data set.

4.3.1 Metric Scoring Thresholds in Mainstem Sites

We evaluated the ten standard B-IBI metrics for mainstem sites to see if the same scoring could be used in both small streams and large, regulated rivers. Table 19 shows the range of values for each of the ten B-IBI metrics for mainstem and tributary sites. For seven of the ten metrics, the range of values was very similar in both mainstem and tributary sites but for three metrics (percent predators, percent tolerant and intolerant taxa richness) the range was smaller in the mainstem sites. We performed t-tests (two-tailed) on all ten metrics to determine whether the mean metric values for the tributary sites differed significantly from the mainstem sites (Table 20). The mean value for both percent predators and intolerant taxa richness in mainstem sites was significantly lower than in the tributary sites (Table 20). In addition, intolerant taxa richness was lower in mainstem reference sites than in tributary reference sites (Table 21).

Table 19. Data ranges for B-IBI metrics in mainstem and tributary sites.

B-IBI Metric	Data range in mainstem	Data range in tributaries
Total Taxa	16.0-39.3	8.3-38.7
Total Mayfly Taxa	4.3-12.3	1.0-10.7
Total Stonefly Taxa	1.0-7.0	0.0-8.0
Total Caddisfly Taxa	1.0-8.7	1.0-9.7
Total Clinger Taxa	0.0-20.0	0.0-21.9
Long-Lived Taxa	2.0-10.0	0.0-9.0
Intolerant Taxa	0.0-4.0	0.0-11.0
Percent Predators	0.0-27.7	0.0-50.1
Percent Tolerants	0.0-15.3	0.0-50.1
Percent Dominance	43.7-89.9	45.0-86.5

Table 20. T-test for difference between mean metric values in all mainstem vs. all tributary sites.

Metric	Mean value in mainstem sites	Standard deviation in mainstem sites	Mean value in tributaries	Standard deviation in tributaries	P value (* = significant at 95% confidence level)
Taxa richness	25.4	4.2	26.5	6.0	0.2553
Mayfly taxa richness	6.7	1.0	6.7	2.1	0.7352
Stonefly taxa richness	3.3	1.6	3.9	2.0	0.0863
Trichoptera taxa richness	5.0	3.1	5.2	1.8	0.5993
Clinger taxa richness	9.5	1.6	9.4	6.2	0.8970
Long-lived taxa richness	5.0	1.0	4.2	2.2	0.2292
Intolerant taxa richness	0.5	0	3.2	3.0	< 0.0001*
Percent Predators	9.5	1.6	11.1	8.7	0.0481*
Percent Tolerant	5.0	7.5	6.8	9.4	0.3118
Percent Dominance	64.3	4.4	62.7	9.9	0.3827

Table 21. Results of t-test between the metric values in mainstem and tributary reference sites.

Metric	Mean value in mainstem reference sites	Standard deviation in mainstem reference sites	Mean value in tributary reference sites	Standard deviation in tributary reference sites	P value (* = significant at 95% confidence level)
Taxa richness	32.1	5.0	32.1	4.7	0.4526
Mayfly taxa richness	8.6	2.4	7.6	1.5	0.2338
Stonefly taxa richness	5.3	0.9	5.2	1.3	0.8567
Trichoptera taxa richness	5.4	1.3	6.8	1.8	0.0915
Clinger taxa	13.5	6.8	10.4	7.1	0.3318

richness					
Long-lived taxa richness	6.5	2.2	5.5	2.2	0.3048
Intolerant taxa richness	1.0	1.3	5.2	3.2	0.0025*
Percent Predators	9.3	5.8	12.8	6.6	0.2386
Percent Tolerant	7.8	4.7	3.9	5.8	0.1231
Percent Dominance	58.0	11.4	59.1	11.3	0.8196

4.3.2 Metrics in Regulated Rivers vs. Unregulated Rivers

We wanted to see if the differences in metric values for percent predators and intolerant taxa richness could be attributed to natural differences between large rivers and small streams or whether the differences in these two metric values could be due to the effects of regulation. In order to test whether differences in metric values are due to regulation and not stream size, we evaluated the values for all metrics in small, regulated rivers and small, non-regulated rivers (Table 22). To minimize the effects of other human land uses and isolate the effects from regulation, we selected sites that have mostly forested watersheds, low human population density and no agricultural land use (Figure 11).

The results of t-tests for differences between the mean metric values in mainstem and tributary sites are presented in Table 23. The percentage of tolerant individuals was significantly higher in the regulated sites ($p<0.05$; $p=.0480$). In addition, the mean number of intolerant taxa was significantly higher ($p=0.0572$) in the free-flowing sites than in the regulated sites. This indicates that the lower value for intolerant taxa richness in mainstem sites could be due to the effects of river regulation rather than to natural variation between mainstem rivers and tributaries, ~~thus~~ supporting our decision to use the same scoring thresholds for mainstem rivers and small tributaries.

Table 22. Regulated and non-regulated small rivers.

Site	Regulated?	Subbasin	Watershed Size (km^2)	population density (# people/km^2)
Lower Bumping River	Yes	Naches	18.1	0
American River	No	Naches	15.0	0
Tieton	Yes	Naches	62.2	2
Little Naches	No	Naches	35.7	0
Cle Elum below dam	Yes	Yakima	39.8	0
Teanaway River	No	Yakima	45.9	4

Table 23. Results of t-tests (unpaired, two-tail) between mean metric values in regulated vs. non-regulated sites

Metric	Hypothesized effect with regulation	Mean in regulated sites	Mean in non-regulated sites	P value (mean)
Taxa richness	Decrease	24.7	25.7	0.7749
Mayfly taxa richness	Decrease	6.7	5.8	0.4229
Stonefly taxa richness	Decrease	4.1	3.7	0.6683
Trichoptera taxa richness	Decrease	4.7	4.7	0.9959
Clinger taxa richness	Decrease	12.2	9.4	0.4421
Long-lived taxa richness	Decrease	4.8	3.4	0.2908
Intolerant taxa richness	Decrease	0.8	3.4	0.0572#
Percent Predators	Decrease	6.6	3.6	0.1909
Percent Tolerant	Increase	6.0	1.0	0.0480*
Percent Dominance	Increase	69.8	61.7	0.2771

*** = significant at 95% confidence level**

 # = significant at 90% confidence level

4.3.3 Road Density and Mainstem Site Condition

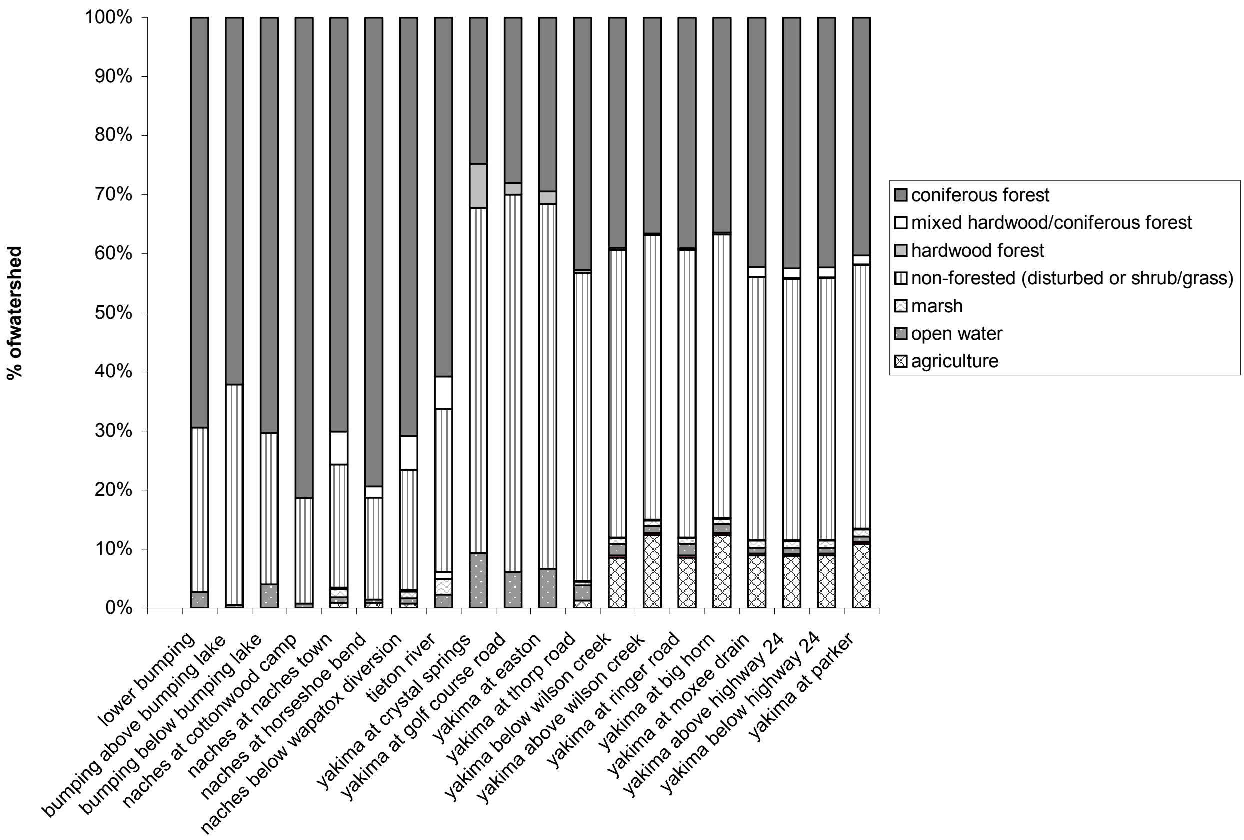
Road densities within 100 and 200 meters of both banks of mainstem sites provided no signal of impact on site condition as in the case of tributary sites (data not shown). This is not surprising in view of the large scale of possible influences of site condition of large rivers and the dominating influence of river regulation on most mainstem sites.

4.3.4 Landcover in Mainstem Sites

The proportion of each type of land cover within the study watersheds is shown in Figure 18. Forest cover for mainstem sites in the Naches subbasin ranged from 62% (Bumping above Bumping Lake) to 89% (Naches below Wapatox diversion), while mainstem Yakima subbasin (mainstem Yakima River and Cle Elum River) sites had watershed forest cover ranging from 29% (Cle Elum) to 43% (Yakima at Thorp Road.) The median watershed forest cover was significantly higher in the Naches subbasin than in the Yakima subbasin (Mann Whitney U-test, $p<0.0001$).

Agricultural land cover was less than 1% of basin cover in all the Naches watersheds, and ranged from 1% (Yakima at Thorp Road) to 12% (Yakima above Wilson Creek and Yakima at Moxee) in the Yakima subbasin. The proportion of developed land cover was too low to be detected for any of the study basins, including for sites that lie within the city of Yakima, and so was discarded from further analysis.

Figure 18. Land cover percentages and distributions across mainstem watersheds. The eight sites on the left side of the chart lie in the Naches subbasin. The remaining thirteen sites are in the mainstem Yakima subbasin.



Land cover classified as “non-forested (generally meaning it was either disturbed or consisted of shrub/grass vegetation) made up 18% to 37% of watershed cover in the Naches subbasin watersheds and 44% to 70% of cover in the Yakima subbasin watersheds. This rather broad category of land cover should be examined in more detail to discern the extent to which the non-forested land is made up of natural vegetation cover; e.g., what proportion of non-forested land cover was classified as disturbed, such as deforested from logging or fire, and what proportion was classified as shrub/grass, a natural vegetation community dominant in the Columbia Plateau ecoregion.

4.3.5 B-IBI in Mainstem Sites

B-IBI scores at mainstem sites ranged from 44 at Lower Bumping River (BUM1) in 2002 to 22 at Tieton River (TIE1) in 2000 (Table 18). Mean B-IBI for all mainstem sites was 30.1, with a mean B-IBI of 32.9 for sites in the Naches subbasin and 28.0 for sites on the mainstem Yakima River. Sites on the upper mainstem Yakima and the Cle Elum River (above and including Yakima at Ringer Road) had an average B-IBI of 26 while sites in the lower Yakima subbasin (Yakima at Big Horn and below) had an average B-IBI of 31.

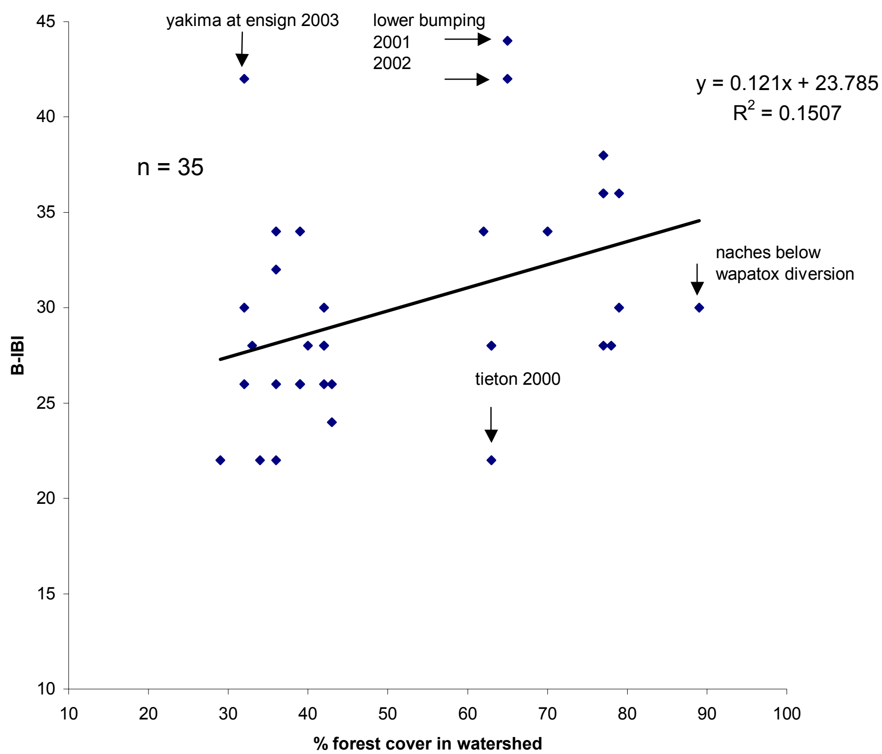
4.3.6 B-IBI vs. Percent Forest Cover in Mainstem Sites

Table 24 summarizes the correlations between percent forest cover from the GAP data and site B-IBI score and the individual metrics for all tributary sites. Figure 19 shows the plot of the linear regression of B-IBI site score against percent forest cover for tributary site-years for which cover data was available.

Table 24. B-IBI and metrics in relation to watershed forest cover for mainstem sites. Shown are R^2 values obtained from linear regression. * = p<0.05, ** = p<=0.01, *** = p<0.001, **** = p<0.0001, ns = not significant.

Index	watershed forest cover (R^2)
B-IBI Score	0.15*
total taxa	0.10^{ns}
ephemeroptera taxa	0.09^{ns}
plecoptera taxa	0.00^{ns}
trichoptera taxa	0.05^{ns}
long-lived taxa	0.31***
intolerant taxa	0.05^{ns}
percent predators	0.01^{ns}
clinger taxa	0.08^{ns}
percent tolerant	0.02^{ns}
percent dominant	0.01^{ns}

Figure 19. B-IBI in relation to watershed forest cover in mainstem sites. Slope = 0.121, R^2 = 0.151.

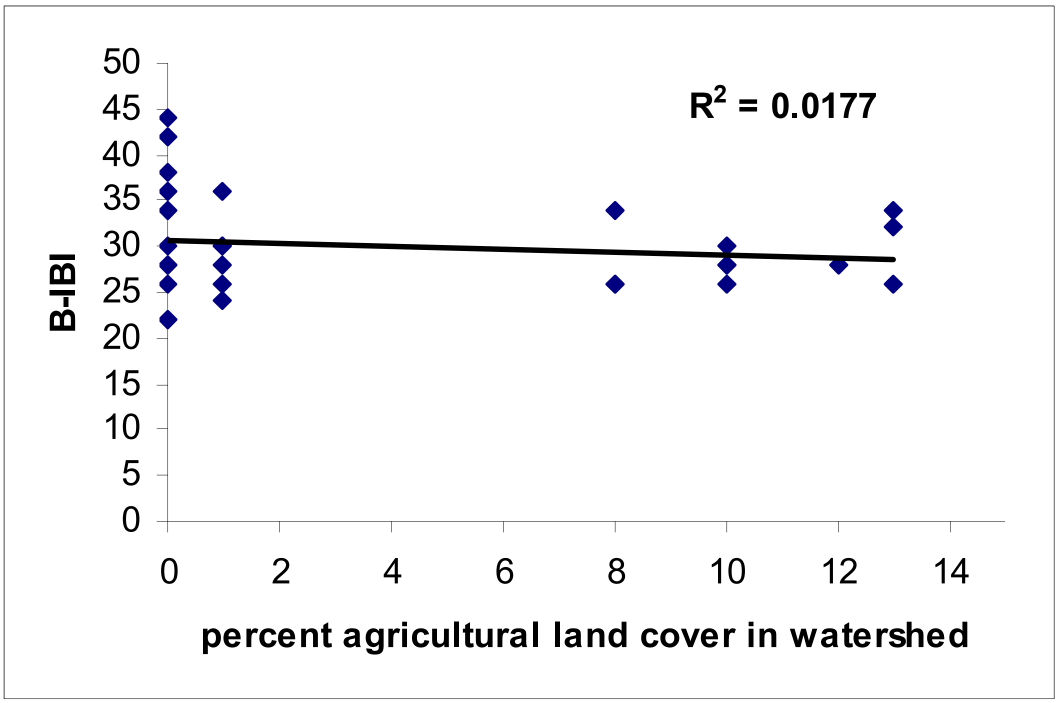


B-IBI in mainstem sites increased as watershed forest cover increased ($p<0.05$; Table 24, Figure 18). The linear regression of site-year score on percent forest cover was significant at an α value of 0.05 ($p<0.05$), with a positive slope of 0.121 and an r-squared value of 0.151. This shows that B-IBI in mainstem rivers responds modestly to percent forest cover: approximately 15 percent of the variance in site score may be accounted for by percent forest cover.

4.3.7 Percent Agricultural Land Cover and Mainstem Site Condition

Figure 20 shows the results of the linear regression of percent of agricultural land cover in mainstem subbasins and tributary site B-IBI score. The regression is not statistically significant ($R^2 = 0.02$. $p> 0.10$).

Figure 20. Percent agricultural land cover and mainstem site B-IBI score.



5.0 DISCUSSION

We employed semi-quantitative, narrative criteria to initially classify mainstem and tributary sites to one of four conditions ranging from severely degraded to reference condition (undisturbed or nearly-undisturbed). The ensuing process of metric testing resulted in a final set of ten (10) metrics and associated scoring criteria for classifying site condition on the basis of B-IBI score (Tables 9 and 13). The final B-IBI (metrics, scoring criteria, and classification criteria) classify tributary and mainstem sites into one of four classes of condition with sufficient accuracy for purposes of assessment and monitoring tributary and mainstem sites.

The final set of metrics and scoring criteria resulted in a substantial revision of the initial (narrative-based) classification of tributary sites (Tables A2 – A4). In general, a majority of sites at elevations above 3500 feet that were initially categorized as reference (condition-1) sites were categorized as condition-2. In a majority of these cases, this result appears to be due to inherent differences in the biological potential of the benthic invertebrate community of higher elevation streams of second and third order. However, of the six sites at elevations at or above 3500 feet accounting for 14 site-years out of the total of 28 site-years of sampling at sites initially categorized as condition-1, only one site accounting for three site-years of sampling failed to achieve a B-IBI score in the condition-1 range. Four of the six sites scored in the condition-1 range in one of two years of sampling. The sixth site scored in the condition-1 range in all three years of sampling (Table A2). Based on these results, we retained the scoring criteria and accepted the downward revision of the condition classification of some of the high elevation sites, rather than create separate scoring criteria for high elevation tributaries.

The remainder of the revisions to the initial classification of site condition consisted largely of one-level upgrades of sites initially classified as condition 2 or 3 (Tables A3 and A4). The result of the revisions for tributary sites with two or more years of sampling was a net reduction in the number of condition-1 sites from ten to seven, a net increase in the number of condition-2 sites from five to twelve, a net reduction of condition-3 sites from six to four, and a net reduction of condition-4 sites from three to one.

Apart from the special case of high (>= 3500 feet) elevation sites, the B-IBI revealed that tributary sites were in better biological condition than they appeared to us upon initial inspection. We believe that this is a perfectly reasonable result. An objective measure of biological condition should result in some adjustment and refinement of our impression of aquatic habitat condition. If we could "tell just by looking" we would have no need of objective, measurable criteria.

There remains the fact that tributary sites exhibited greater levels of interannual variation in B-IBI score than expected from other published B-IBI's. There may be four reasons for this result. First, other B-IBI studies have either involved fewer years of sampling than our study, or the number of sites that were sampled in two or more years have made up a smaller proportion of the total sampling effort than in our study. We may therefore have had a higher probability of identifying interannual variation in site scores because we sampled a greater proportion of our sites for two or more years. Second, we may have sampled inconsistently at sites during different years. Third, our metrics and our scoring criteria may be faulty and fail to properly reflect the biological condition of sites. Fourth, sites in the Yakima may be inherently more variable when measured with B-IBI metrics than other regions where B-IBI has been studied.

We think that there may be some truth to the first explanation but have not investigated the matter in the detail required. Other published reports of B-IBI have not provided the level of detail on interannual variation of site score that we have in the present report. For example, the major published analysis of the performance of B-IBI using data from three consecutive years in the Umpqua National Forest (Fore et al 1996) did not evaluate inter-annual variability in B-IBI score at individual sites (but see Fore, 2003). So this possibility remains to be investigated.

We do not think that our sampling was inconsistent at sites between years. All five individuals who conducted the field sampling over the four years were well trained. Three of the five had extensive prior experience in B-IBI sampling. The other two were extensively trained by one of us and initially assisted one or the other of the three experienced individuals for two to three years before sampling on their own in year four. There was, therefore, a high degree of proficiency and consistency between individuals in the conduct of the sampling protocols.

We do not believe that the metrics are at fault. The same ten metrics that have been validated for the Puget Sound Lowlands and the Clackamas River basin were chosen for the Yakima B-IBI.. Interannual variation in site scores in these regions has not been noted to be a problem. Moreover, mainstem sites displayed noticeably less interannual variation than tributary sites using the same set of metrics and scoring criteria. At sites on regulated segments of mainstem rivers, this is likely a reflection of the pervasive and generally homogenizing influence of river regulation. If the metrics were faulty we would expect regulated mainstem sites to show a pattern of interannual variation more similar to that of tributaries, rather than reflect the influence of regulation as they appear to do. In addition, unregulated mainstem sites also showed no greater interannual variation than regulated rivers and no significant difference from comparable regulated sites in the majority of individual metrics (Table 23).

Consequently, we are lead to conclude that tributary sites may be more variable in the Yakima basin than west of the Cascades. However, the degree of interannual variation appears to be of sufficiently low magnitude to enable the proposed B-IBI to provide a sufficiently accurate classification of site condition for the purposes of assessment and monitoring for which it is intended. We noted in the sub section *4.1.1 Metrics and Metric Scoring Criteria for Tributaries* in Results that a standard deviation in B-IBI score of 6.0 would produce 50% confidence limits of +/- 4.0. Table 25 lists all site-year scores for tributary sites with two to four years of sampling. Sites-codes highlighted in bold font indicate sites that received different condition classifications in different years but in which the absolute value of none of the scores differed by more than 4.0 from one of the three classification threshold scores (20, 28, or 38). Sites-codes and scores highlighted in italicized bold font indicate sites that received different condition classifications in different years but in which the absolute value of one or more scores differed by greater than 4.0 from a classification threshold. There are four sites accounting for eight site-years that fall in this latter category, NTEA (North Fork Teanaway), OAK1 (Oak Creek), TEA1 (mainstem Teanaway) and WEN1 (North Fork Wenas Creek).

Table 25. Tributary B-IBI scores and condition classification by site-year. Site-codes in bold indicate sites at which between-year scores result in different condition classification but at which no score differs by more than 4.0 from one of the classification threshold scores (20, 28, or 38). Site-codes and scores in bold-italics indicate sites in which between-year scores differ by more than 4.0 from a classification threshold.

Site-Year	B-IBI Score	Condition
AHT1.00	34	2
AHT1.02	38	1
AHT2.00	42	1
AHT2.01	38	1
AHT2.02	48	1
AHT3.00	30	2
AHT3.01	22	3
AHT3.03	24	3
AME2.01	30	2
AME2.02	32	2
BIG1.00	34	2
BIG1.01	34	2
BIG1.02	36	2
CAB1.01	42	1
CAB1.02	34	2
CAB1.03	38	1
CHE.00	28	2
CHE.03	22	3
COO1.00	40	1
COO1.01	44	1
COO1.02	46	1
COW1.01	26	3
COW1.02	24	3
COW1.03	22	3
MAN1.00	38	1
MAN1.02	36	2
MAN2.00	36	2
MAN2.01	44	1
NAN1.00	34	2
NAN1.02	42	1
NIL1.00	36	2
NIL1.02	38	1
NTA1.00	32	2
NTA1.02	42	1

NTA2.01	40	1
NTA2.02	46	1
NTEA.00	***42***	1
NTEA.02	***28***	2
OAK1.00	***40***	1
OAK1.02	***30***	2
RAT1.00	34	2
RAT1.01	34	2
RAT1.02	38	1
RAT1.03	36	2
TAN2.01	42	1
TAN2.02	36	2
TAN3.01	28	2
TAN3.03	28	2
TEA2.01	***32***	2
TEA2.02	***38***	1
WEN1.00	***26***	3
WEN1.03	***36***	2
WEN2.01	18	4
WEN2.02	16	4
WEN2.03	20	3
WID1.00	24	3
WIDE.03	20	3

North Fork Wenas Creek, for example, scored 26 in 2000 which is two points below the threshold of 28 separating condition-2 sites from condition-3 sites, and scored 36 in 2003 which is eight points above the threshold. By contrast, TAN2 (Taneum Creek on Bureau of Reclamation property) scored 42 in 2001 (four points above the threshold value of 38) and 36 in 2002 (two below the threshold). There are eight other sites accounting for 18 site-years in this second category.

Assuming that no change has actually occurred at the TAN2 site between 2001 and 2002, the two scores at this site indicates there is better than 50% probability that the true condition of the site is close to the border between condition-1 and condition-2. In the other four cases, there is much less confidence regarding which of the two conditions the site is really in. In the remaining eleven sites (29 site-years), site scores in multiple years are either all within a single condition class or are close to a threshold value with site-year scores differing by four points or less from the threshold value.

This indicates that three years of sampling will generally provide an accurate indication of site-condition and variability in site-condition. Four years of sampling will resolve particularly anomalous cases, as in the case of the mainstem of the Naches River at Cottonwood campground discussed in the Results section. These considerations apply to mainstem sampling as well which displayed less interannual variation than tributary sites (Tables 17 and 18).

The magnitude of interannual variability of tributary sites and the possible causes for it will become better resolved as more data accumulates. Like all other biomonitoring tools, the B-IBI will continue to be refined and improved the more it is employed in individual regions.

We evaluated the response of site condition (B-IBI score) to large landscape (watershed scale) variables related to human disturbance, road density, percent forest cover, percent agricultural land cover, and percent developed land cover. Only percent forest cover and percent agricultural land cover (tributaries only) showed a relationship to B-IBI score. In each of these cases the relationship was in the expected direction (B-IBI increasing with increasing forest cover and decreasing with increasing agricultural land cover, figures 15, 16, and 19).

The B-IBI revealed that tributary site conditions are moderately sensitive to the percentage of forest cover. The regression of B-IBI score against percentage forest cover in tributary watersheds had a positive regression slope (0.141) with an R^2 value of 0.28 and a p-value <0.001. Mainstem sites also displayed sensitivity to percent forest cover, with the slope of the regression only slightly smaller than for tributaries (0.121 but with smaller R^2 and p-value (0.151, $p<0.05$).

The regression of B-IBI score against percent agricultural land cover in tributary watersheds had a statistically significant negative relationships (slope = -0.1813, R^2 = 0.1828, $p<0.0001$). The influence of agricultural land cover is complex, however, and depends upon finer-scale features than we were able to analyze. For example, the lowest site scores occurred at sites with 10 to 30 percent agricultural land cover, Sites with nearly 100 percent agricultural land cover were able to score slightly higher, though no site with greater than 1 percent agricultural cover scored higher than 32. Scores of sites with less than 1 percent agricultural land cover ranged from 22 to 48 (Figure 16).

A potential limitation associated with using the WDFW land cover data as a measure of human disturbance lies in the different goals for the GAP study and for our study. The primary goal for the GAP land cover data was to map vegetation land cover, not to map human modified land cover, i.e. developed land or agricultural land. Therefore these non-natural land cover types may not have been mapped at the same level of accuracy as natural vegetation types.

Another limitation is the size of the polygons delineated in the GAP data set. While the minimum mapping size used was apparently 0.01 square kilometers, the average polygon size for the data used in our study was 1.15 square kilometers. For agricultural lands, the average polygon size was even higher at 2.63 square kilometers. Polygons for developed lands were mapped at a higher resolution (average of 0.47 square kilometers) than agricultural lands but still it is unclear whether these land cover polygons are small enough to sufficiently capture human influence, particularly when human disturbance occurs in small-scale, patchy distributions as seems likely in rural, developing watersheds.

In contrast to percent forest cover, B-IBI score showed no clear relationship to watershed road density within either 100 or 200 meters of either side of the stream bank. Our inability to detect any influence of road density at these scales may be due in part to the quality of the GAP road data. We were unable, for example, to estimate road crossing or to determine road condition. There was also uncertainty regarding the precision of the overlay of the road polygons to the stream layers, which would have affected the accuracy of the delineations of the buffers.

The results may, however, be due to a combination of relatively low road densities in most of the basins and the condition of riparian buffers within 100 meters of stream banks and not merely to the quality of the GAP data. In addition, we suspect that the catchment scale may be too large to detect road impacts at the channel unit (riffle) scale.

Road density and condition measured at a finer spatial scale relative to sample sites is likely required to assess road impacts.

6.0 CONCLUSIONS AND RECOMMENDATIONS

6.0.1 Recommended Application of the Index for Monitoring and Assessment

We recommend that the index be employed in tributaries and mainstem sites of the Yakima River upstream of Union Gap, including the Naches River, as a tool for the assessment of site condition, monitoring of impacts, and monitoring of site response to restoration actions. We recommend that site assessment and monitoring at specific sites be undertaken for periods of two or more consecutive years, wherever budget allows. This is particularly important at sites that it is believed will exhibit intermediate levels of disturbance/impairment (condition-2 and condition-3). The data shows that condition-4 (severely impaired) and condition-1 (minimally-disturbed) sites display little interannual variation compared to condition-2 and condition-3 sites. In order to obtain reliable results for monitoring and assessment at these kinds of sites using the proposed B-IBI two to four consecutive years of data collection are required. Although informative results can be obtained from one year of data using B-IBI, we are distrustful of snapshot monitoring and assessment of aquatic habitat. This is less a concern with follow-up or spot-check monitoring of a site that has recently been sampled for one or more years. Single-year check-up monitoring of previously sampled sites can be very informative using B-IBI. The present study has resulted in a significant baseline of sampled sites and B-IBI data for this purpose. For sites that have not been previously sampled, we recommend that if it is important enough to monitor a site at all, it is worth allocating enough resources to obtain several years of data.

B-IBI is particularly cost-effective in this regard. Sampling costs are modest. No more than two (2) hours are required to collect the required number of invertebrate samples and to take basic measurements of stream channel and riparian area condition. Processing of the three replicate samples that are required to be taken at each site, including sorting and identification requires an average of 20 hours of laboratory time. The total cost for these activities per sample site and sampling occasion is approximately $600.00.

6.0.2 Unresolved Issues and Recommended Future Research

We also recommend that B-IBI be used as a monitoring tool in conjunction with fish habitat restoration projects and fish habitat preservation projects. It is important to emphasize that the B-IBI directly measures the biological condition of the benthic food web and indirectly the biophysical condition of a stream/river site. It does not measure the ability of a site or associated stream reach to support fish taxa of interest such as salmonids. For purposes of monitoring the effect of stream habitat restoration activities on salmonid species, the B-IBI is best employed in conjunction with other measures of stream habitat condition such as physical conditions in the stream channel and the associated riparian area and indices of fish population condition and individual fish condition.

The B-IBI has yet to be fully evaluated in the context of salmonid fish use, individual fish condition, and abundance. This can only be accomplished by employing the B-IBI in tandem with fish population monitoring activities, both those activities associated with fish populations in healthy stream habitats and those intended to restore fish habitat quality. This should be an important area of future research because the B-IBI has potential to provide a direct assessment of the quality of a stream reach from a fish's perspective because it measures both the condition of the benthic food web and the condition of specific kinds of invertebrate taxa important to salmonid fish.

The specific link between individual metrics employed in a B-IBI and the feeding ecology of salmonid juveniles has yet to be fully taken advantage of. It would be particularly valuable to employ the B-IBI in conjunction with a study of the functional significance of benthic invertebrate taxa for drift feeding salmonids as exemplified in Rader's "functional classification of the drift" (Rader, 1997). Such a study could provide important information on the correlation between B-IBI index scores and the conditions at the stream channel and stream reach scale that are significant from the point of view of salmonid feeding ecology. We recommend that such a study be developed and funded in the Yakima basin.

7.0 REFERENCES

Adams, J.W. 2001. “Using aquatic macroinvertebrates to assess stream condition in the Clackamas River Basin, OR.” Master’s Thesis. University of Washington. 72 pp.

Doberstein, Craig P., J. R. Karr, and L. L. Conquest. 2000. “The effect of fixed-count subsampling on macroinvertebrate biomonitoring in small streams”. Freshwater Biology, vol. 44: 355-371.

Fore, Leska S., 2003. “Developing biological indicators: lessons learned from mid-Atlantic streams.” EPA 903/R-003/003. U.S. Environmental Protection Agency, Office of Environmental Information and Mid-Atlantic Integrated Assessment Program, Region 3, Ft. Meade, Md.

Fore, Leska S., J. R. Karr, and L. L. Conquest. 1994. “Statistical properties of an index of biological integrity used to evaluate water resources”. Canadian Journal of Fisheries and Aquatic Sciences, Vol. 51, 1077 – 1087.

Fore, Leska S., J. R. Karr, and R. W. Wisseman. 1996. “Assessing invertebrate responses to human activities: evaluating alternative approaches”. Journal of the North American Benthological Society, vol. 15: 212 – 231.

Karr, James R. and E. Chu. 1999. <u>Restoring Life to Running Waters.</u> Island Press, Washington D. C.

Karr, James R. 1998. "Rivers as sentinels: using the biology of rivers to guide landscape management". In Naiman, R. J. and B. E. Bilby (eds.) <u>River Ecology and Management</u>, Springer-Verlag, New York, N. Y.

Kleindl, W. J. 1995. "A benthic index of biotic integrity for Puget Sound lowland streams, Washington, USA." MS thesis, University of Washington, Seattle.

Morley, S. A. 2000. "Effects of urbanization on the biological integrity of Puget Sound lowland streams: restoration with a biological focus." MS thesis, University of Washington, Seattle.

Patterson, A. J. 1996. "The effect of recreation on biotic integrity of small streams in Grand Teton National Park." MS thesis, University of Washington, Seattle.

Rader, Russel B . 1997. "A functional classification of the drift: traits that influence invertebrate availability to salmonids". Canadian Journal of Fisheries and Aquatic Sciences 54: 1211 – 1234.

Stanford, Jack A. and J. V. Ward. 1992. "Management of aquatic resources in large catchments: recognizing interactions between ecosystem connectivity and environmental disturbance." In Naiman, R. J. (ed.), <u>Watershed Management</u>. Springer-Verlag, New York. Pp. 91 – 124.

Stanford, Jack A., and J. V. Ward. 1995. "Ecological connectivity in alluvial river ecosystems and its disruption by flow regulation". Regulated Rivers Research & Management, vol. 11, 105 – 119.

Stanford Jack A., J. V. Ward, William J. Liss, Christopher Frissell, Richard N. Williams, James A. Lichatowich, and Charles C. Coutant. 1996. "A protocol for restoration of regulated rivers". Regulated Rivers Research & Management, vol. 12, 391 - 413.

8.0 APPENDICES

Table A1. Tributary and Mainstem River Sample Site Locations.

	Site Code	Sept. 3, 2004 Site Description	Elevation (ft.)	Subbasin	Latitude	Longitude	Ecoregion	B-IBI Score	Condition Category	No.Years Sampled
		Tributary Sites								
1	AHT1	North Fork Ahtanum	4700	AHTANUM	46 31.51	121 09.81	East Cascades	36	2	2
2	AHT2	South Fork Ahtanum	2500	AHTANUM	46 30.36	120 55.54	East Cascades	43	1	3
3	AHT3	Lower Ahtanum	1000	AHTANUM	46 32.313	120 28.790	Columbia Plateau	25	3	3
4	AME1	American at Hell's Crossing C.G.	2900	AMERICAN	46 57.911	121 15.936	Cascades	32	2	1
5	AME2	American above Hell's Crossing	3300	AMERICAN	N 46 56.920	W 121 18.057	Cascades	31	2	2
6	AME3	American at Americdan Forks C.G.	2700	AMERICAN	46 58.642	121 09.599	Cascades	38	1	1
7	BIG1	Big Creek	3600	BIG	47 09.049	121 14.923	North Cascades	35	2	3
8	BIG2	Big Creek 2	2200	BIG	ND	ND		38	1	1
9	CAB1	Cabin Creek	2300	CABIN	47 14.053	121 13.612	North Cascades	38	1	3
10	CHE1	Cherry Creek	1500	CHERRY	46 57.404	120 28.621	Columbia Plateau	25	3	2
11	COO1	Cooke Creek	3500	COOKE	47 09.411	120 20.530	East Cascades	43	1	3
12	COW1	Cowiche Creek	1500	COWICHE	46 37.796	120 39.760	Columbia Plateau	24	3	3
13	LNA1	M. F. Little Naches	3100	LITTLE NACHES	47 04.75	121 15.32	Cascades	36	2	1
14	LNA2	Little Naches above junction with Bumping R.	2600	LITTLE NACHES	N 46 59.239	W 121 05.584		38	1	1
15	MAN1	Manastash upstream	4300	MANASTASH	47 02.221	120 57.260	East Cascades	37	2	2
16	MAN2	Manastash downstream	2800	MANASTASH	46 58.06	120 48.29	East Cascades	40	1	2
17	MOX1	Moxee Drain	1000	MOXEE	46 32.173	120 27.253	Columbia Plateau	14	4	1
18	NAN1	Naneum 1	3500	NANEUM	47 13.235	120 26.319	East Cascades	38	1	2
19	NAN2	Naneum below Charleton Rd. Br.	2600	NANEUM	47 06.176	120 28.530	Columbia Plateau	30	2	1
20	NAN3	Naneum 3	2800	NANEUM	47 08.184	120 28.475	Columbia Plateau	42	1	1
21	NIL1	Nile Creek	2400	NILE	46 51.055	121 00.808	East Cascades	37	2	2
22	NTA1	N. F. Taneum above FR 4501Crossing	3700	TANEUM	47 07.515	121 03.957	North Cascades	37	2	2
23	NTA2	N. F. Taneum below FR 4501Crossing	3700	TANEUM	47 07.495	121 03.913	North Cascades	43	1	2
24	OAK1	Oak Creek	2500	OAK	46 43.896	120 53.703	East Cascades	35	2	2
25	RAT1	Rattlesnake above N.F. confluence	2700	RATTLESNAKE	46 48.570	121 04.148	East Cascades	36	2	4
26	RAT2	Rattlesnake lower	2000	RATTLESNAKE	46 48.978	120 56.614	East Cascades	30	2	1
27	SBE1.01	Aschbach Springbrook #3	1300	NACHES	N 46 40.18	W 120 38.90	Columbia Plateau	24	3	1
28	TAN1	Taneum at Taneum C.G.	2700	TANEUM	47 06.413	120 51.425	Columbia Plateau	30	2	1
29	TAN2	Taneum @ BOR/Rocky Mt.Elk Fdn. Prop.	2000	TANEUM	47 04.968	120 45.130	Columbia Plateau	39	1	2
30	TAN3	Taneum below I-90 Br.	1900	TANEUM	47 04.968	120 43.893	Columbia Plateau	28	2	2
31	TAN4	Taneum below confluence of S. Branch Canal	2000	TANEUM	47 04.892	120 44.913	Columbia Plateau	30	2	1
32	TEA1	N. F. Teanaway above N. F. Br.	2400	TEANAWAY	47 17.371	120 51.594	North Cascades	35	2	2
33	TEA2	Teanaway mainstem (RR Br.)	1900	TEANAWAY	47 12.234	120 46.881	North Cascades	35	2	2
34	UMT1	Umtanum	2600	UMTANUM	46 53.966	120 38.583	Columbia Plateau	26	3	1
35	WEN1	N. F. Wenas	2500	WENAS	46 53.799	120 47.825	East Cascades	31	2	2
36	WEN2	Wenas @ BOR Property	1200	WENAS	N 46 41.80	W 120 29.69	Columbia Plateau	18	4	3
37	WID1	Wide Hollow in Union Gap	1000	WIDE HOLLOW	46 32.569	120 28.531	Columbia Plateau	22	3	2
										70

Table A1. Tributary and Mainstem River Sample Site Locations.

	Site Code	Sept. 3, 2004 Site Description	Elevation (ft.)	Subbasin	Latitude	Longitude	Ecoregion	B-IBI Score	Condition Category	No.Years Sampled
		Mainstem Sites								
1	BUM1	Bumping R. at Cedar Springs C.G.	2800	BUMPING	N 46 58.254	W 121 03.747	East Cascades	43	1	2
2	BUM2	Bumping River above Bumping Reservoir	3700	BUMPING	46 49.787	121 22.561	Cascades	34	2	1
3	BUM3	Bumping River below Bumping Crossing Br.	3400	BUMPING	46 52.874	121 16.784	Cascades	34	2	1
4	CLE1	Cle Elum at Salmon la Sac	2400	CLE ELUM	47 23.472	121 05.779	North Cascades	22	3	1
5	CLE2	Cle Elum above Lake	2300	CLE ELUM	47 20.908	121 06.735	North Cascades	22	3	1
6	CLE3	Cle Elum at Roslyn	2000	CLE ELUM	N 47.11	W 121 00.86	North Cascades	22	3	1
7	NAC1	Naches at Cottonwood CG	2200	NACHES	46 54.38	121 01.56	East Cascades	35	2	4
8	NAC2	Naches at Naches	1400	NACHES	46 43.49	120 42.19	Columbia Plateau	33	2	2
9	NAC3	Naches at Horseshoe Bend	1700	NACHES	N 46 45.142	W 120 49.083	Columbia Plateau	28	2	1
10	NAC4	Naches at Wapatox	1600	NACHES	N 46 44.878	W 120 46.890	Columbia Plateau	30	2	1
11	TIE1	Tieton	1900	TIETON	46 42.987	120 51.599	East Cascades	26	3	3
12	YAK01 *	Yakima at Golf Course Rd.	2100	UPPER YAKIMA	47 11.168	121 02.640	North Cascades	33	2	1
13	YAK02	Yakima @ RRR	1700	UPPER YAKIMA	47 04.469	120 39.692	North Cascades	25	3	2
14	YAK03	Yakima @ Ringer Rd	1400	UPPER YAKIMA	46 55.416	120 31.077	Columbia Plateau	26	3	2
15	YAK04	Yakima @ Big Horn	1400	UPPER YAKIMA	46 53.574	120 29.438	Columbia Plateau	34	2	2
16	YAK05 **	Yakima @ Moxee Drain	1000	LOWER YAKIMA	46 32.239	120 27.399	Columbia Plateau	26	3	1
17	YAK06	Yakima @ Parker	900	LOWER YAKIMA	46 29.776	120 26.355	Columbia Plateau	28	2	1
18	YAK07 *	Yakima @ WDFW boat launch	2100	UPPER YAKIMA	N 47 12.90	W 121 05.45	North Cascades	34	2	2
19	YAK08	Yakima @ Crystal Springs	2400	UPPER YAKIMA	47 18.386	121 18.802	North Cascades	30	2	1
20	YAK09 ***	Yakima upstream Hwy 24	1000	LOWER YAKIMA	N 46 35.11	W 120 27.61	Columbia Plateau	29	2	2
21	YAK10 **	Yakima downstream Hwy 24	1000	LOWER YAKIMA	N 46 33.33	W 120 27.82	Columbia Plateau	28	2	1
22	YAK11	Yakima Upstream Wilson Creek	1400	UPPER YAKIMA	46 55.056	120 30.605	Columbia Plateau	34	2	1
23	YAK12	Yakima downstream Wilson Creek	1400	UPPER YAKIMA	46 54.607	120 30.499	Columbia Plateau	26	3	1
										35
										105

* : Yak01 and Yak07 are in the same river segment but in different reaches and are considered the same site for purposes of analyzing interannual variaiton in B-IBI scores due to similarity of reach conditions.

** : Yak05 and Yak10 are in the same river segment but in different reaches and are considered the same site for purposes of analyzing interannual variaiton in B-IBI scores due to similarity of reach conditions.

*** : Two distinct sites in the same segment but in different reaches and on opposites banks of the river were sampled but are considered to be the same site for purposes of analyzing interannual variaiton in B-IBI scores due to similarity of reach conditions.

Table A2. Comparison of site-years classified as condition 1 by narrative standards with site classification determined by recommended B-IBI scoring criteria. A negative difference means that the scoring criteria classify the site as being in poorer condition than the narrative classification. A positive difference means the scoring criteria classify the site as being in better condition than the narrative classification.

Site-Year	B-IBI Score	Narrative (N) Condition	Score-Based (S) Condition	Difference (N – S)
AHT1.00	34	1	2	-1
AHT1.02	38	1	1	0
AHT2.00	42	1	1	0
AHT2.01	38	1	1	0
AHT2.02	48	1	1	0
AME2.01	30	1	2	-1
AME2.02	32	1	2	-1
BIG1.00	34	1	2	-1
BIG1.01	34	1	2	-1
BIG1.02	36	1	2	-1
BIG2.02	38	1	1	0
COO1.00	40	1	1	0
COO1.01	44	1	1	0
COO1.02	46	1	1	0
LNA1.00	36	1	2	-1
MAN1.00	38	1	1	0
MAN1.02	36	1	2	-1
NAN1.00	34	1	2	-1
NAN1.02	42	1	1	0
NAN3.02	42	1	1	0
NTA1.00	32	1	2	-1
NTA1.02	42	1	1	0
NTA2.01	40	1	1	0
NTA2.02	46	1	1	0
RAT1.00	34	1	2	-1
RAT1.01	34	1	2	-1
RAT1.02	38	1	1	0
RAT1.03	36	1	2	-1

Table A3. Comparison of site-years classified as condition 2 by narrative standards with site classification determined by recommended B-IBI scoring criteria. A negative difference means that the scoring criteria classify the site as being in poorer condition than the narrative classification. A positive difference means the scoring criteria classify the site as being in better condition than the narrative classification.

Site-Year	B-IBI Score	Narrative (N) Condition.	Score-Based (S) Condition.	Difference (N – S)
AME1.00	32	2	2	0
AME3.02	38	2	1	1
LNA2.01	38	2	1	1
MAN2.00	36	2	2	0
MAN2.01	44	2	1	1
NIL1.00	36	2	2	0
NIL1.02	38	2	1	1
OAK1.00	40	2	1	1
OAK1.02	30	2	2	0
RAT2.00	30	2	2	0
TAN1.00	30	2	2	0
TAN2.01	42	2	1	1
TAN2.02	36	2	2	0
UMT1.00	26	2	3	-1
WEN1.00	26	2	3	-1
WEN1.03	36	2	2	0

Table A5. Score-Based Classification of 24 tributary Sites With 2 to 4 years of sampling data

Site-Code	Score-Based (S) Condition
AHT2	1
CAB1	1
COO1	1
MAN2	1
NAN1	1
NTA2	1
TAN2	1
AHT1	2
AME2	2
BIG1	2
MAN1	2
NIL1	2
NTA1	2
OAK1	2
RAT1	2
TAN3	2
TEA1	2
TEA2	2
WEN1	2
AHT3	3
CHE1	3
COW1	3
WID1	3
WEN2	4

Figure A1.

Figure A2.

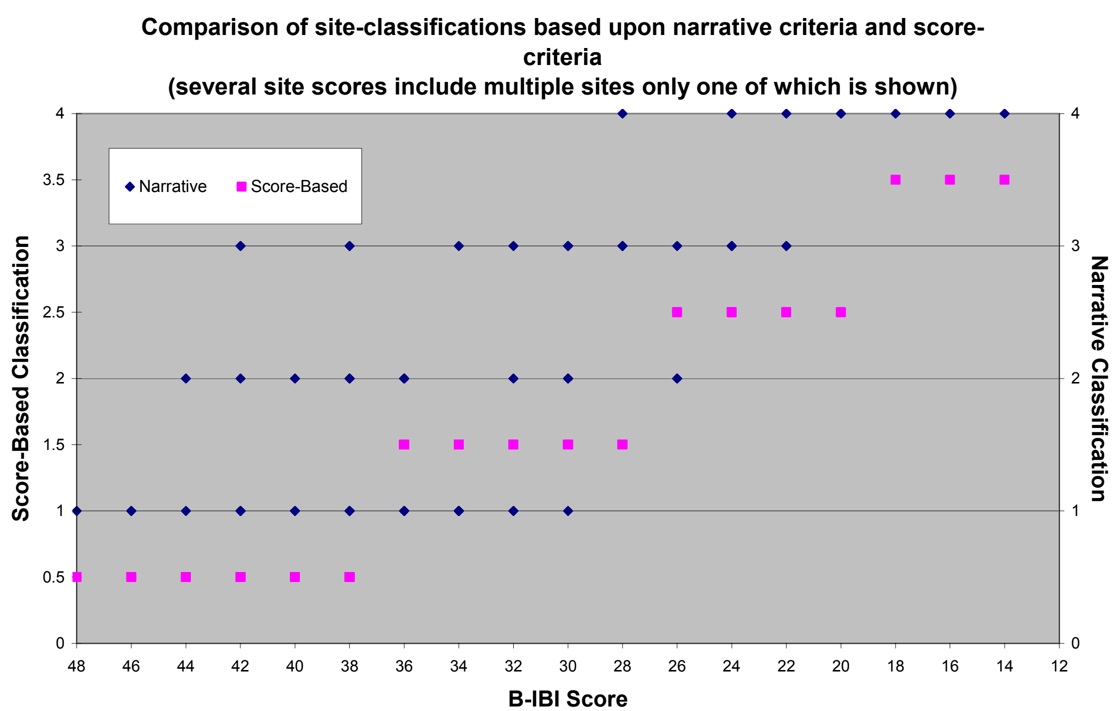


Figure A3.

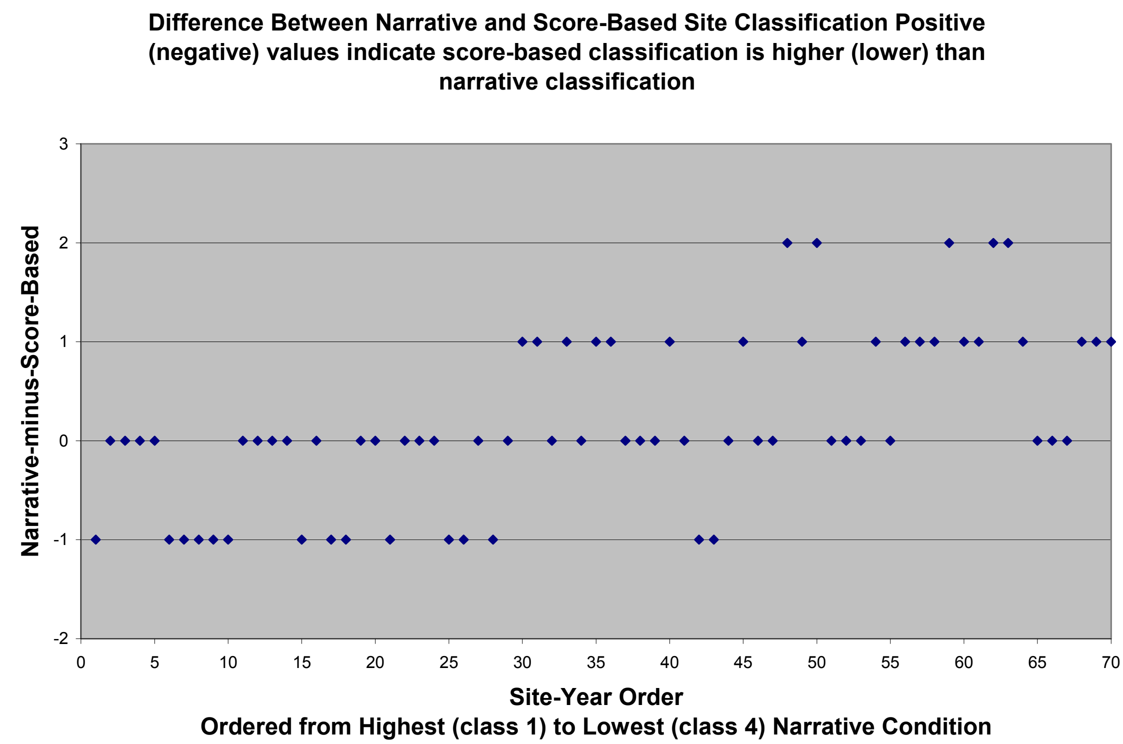


Figure A4